

საქართველოს ტექნიკური უნივერსიტეტი

შალვა მალიძე

ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი
ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა
გაუმჯობესება

წარმოდგენილია დოქტორის აკადემიური ხარისხის მოსაპოვებლად
სადოქტორო პროგრამა „ინფორმატიკა“
შიფრი 0401

საქართველოს ტექნიკური უნივერსიტეტი
თბილისი, 0175, საქართველო
2019 წელი

საავტორო უფლება © 2019 წელი, შალვა მალიძე

საქართველოს ტექნიკური უნივერსიტეტი

ინფორმატიკისა და მართვის სისტემების ფაკულტეტი

ჩვენ, ქვემოთ ხელისმომწერი ვადასტურებთ, რომ გავეცანით შალვა მალიძის მიერ შესრულებულ სადისერტაციო ნაშრომს დასახელებით: „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ და ვაძლევთ რეკომენდაციას საქართველოს ტექნიკური უნივერსიტეტის ინფორმატიკის, მართვისა და ხელსაწყოთმშენებლობის საუნივერსიტეტო სადისერტაციო საბჭოში მის განხილვას დოქტორის აკადემიური ხარისხის მოსაპოვებლად.

-----, ----- 20-- წელი

ხელმძღვანელი: _____

რეცენზენტი: _____

რეცენზენტი: _____

საქართველოს ტექნიკური უნივერსიტეტი

2019

ავტორი: შალვა მალიძე

დასახელება: „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“

სადოქტორო პროგრამა: ინფორმატიკა

ხარისხი: ინჟინერიის დოქტორის აკადემიური ხარისხი ინფორმატიკაში

სხდომა ჩატარდა: _____

ინდივიდუალური პიროვნებების ან ინსტიტუტების მიერ ზემომოყვანილი დასახელების დისერტაციის გაცნობის მიზნით მოთხოვნის შემთხვევაში მისი არაკომერციული მიზნებით კოპირებისა და გავრცელების უფლება მინიჭებული აქვს საქართველოს ტექნიკურ უნივერსიტეტს.

ავტორის ხელმოწერა

ავტორი ინარჩუნებს დანარჩენ საგამომცემლო უფლებებს და არც მთლიანი ნაშრომის და არც მისი ცალკეული კომპონენტების გადაბეჭდვა ან სხვა რაიმე მეთოდით რეპროდუქცია დაუშვებელია ავტორის წერილობითი ნებართვის გარეშე.

ავტორი ირწმუნება, რომ ნაშრომში გამოყენებული საავტორო უფლებებით დაცულ მასალებზე მიღებულია შესაბამისი ნებართვა (გარდა იმ მცირე ზომის ციტატებისა, რომლებიც მოითხოვენ მხოლოდ სპეციფიურ მიმართებას ლიტერატურის ციტირებაში, როგორც ეს მიღებულია სამეცნიერო ნაშრომების შესრულებისას) და ყველა მათგანზე იღებს პასუხისმგებლობას.

წინამდებარე ნაშრომს ვუძღვნი მამის ხსოვნასა და ჩემ ოჯახს - დედას, დებს, მეუღლესა და პატარა ვაჟკაცს. ჩემი დიდი სურვილია უკვე კარს მომდგარ და სანახევროდ უკვე ამოქმედებულ ციფრულ ეპოქაში, ჩემს თომას, რომელიც ჯერ ერთი წლისაც არ არის, საერთო მსოფლიო მოთხოვნების შესაბამისი განათლების მიღების საშუალება ჰქონდეს საქართველოში და თანაც ქართული ენით. გარდა ამისა, ნაშრომი ეძღვნება ჩვენი - ქართველებისა და აფხაზების დიდი წინაპრების ხსოვნას, რადგან ღრმად მწამს, როგორც ქართულ-აფხაზური ერთობის აღდგენისა, ასევე იმისაც, რომ ახლო მომავალში ერთიანი ქართული სახელმწიფოს პირობებში ჩვენ - ქართველები და აფხაზები შევძლებთ ევროკავშირში, უფრო ზოგადად კი, მომავლის კულტურულ სამყაროში ქართული და აფხაზური ენებით შესვლასა და სრულფასოვან თანამონაწილეობას და თანაც ისე, როგორც ამას გვკარნახობს ქართულ-აფხაზური მრავალსაუკუნოვანი კულტურულ - ისტორიული წარსული.

რეზიუმე

ქართული ენის, მით უფრო აფხაზური ენის ენობრივი რესურსებითა და ტექნოლოგიებით მხარდაჭერის მოცულობა და ხარისხი საგანგაშოდ ჩამორჩება იმ ამ მხრივ არასაკმარისად მხარდაჭერილი ევროპული ენებიდან თითქმის ნებისმიერის მოცულობასა და ხარისხს, რომლებიც ევროსტრუქტურების დაფინანსებით მეტა-ქსელის მიერ 2010 – 2012 წლებში განხორციელებული ორწლიანი საერთო ევროპული კვლევის „ევროპის ენების ციფრულ ეპოქაში“ თანახმად ციფრული კვდომის საფრთხის ქვეშ არიან დღეს უკვე კარს მომდგარ და სანახევროდ უკვე ამოქმედებული ციფრულ ეპოქაში. აღნიშნულის გათვალისწინებით, წინამდებარე სადოქტორო კვლევა „ქართული ჰკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“, რომელიც არის საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის გრძელვადიანი პროექტების „ქართული ენის ტექნოლოგიური ანბანი“ და „აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის გეგმა-პროგრამა“ კიდევ ერთი მნიშვნელოვანი ქვეპროექტი, მიმართულია საქართველოს სახელმწიფო ენების - ქართულისა და აფხაზურის სრული ტექნოლოგიური უზრუნველყოფის მიზნების მიღწევაზე. ეს სხვა სიტყვებით, ნიშნავს იმას, რომ სადოქტორო კვლევა მიმართულია ქართული და აფხაზური ენების ციფრული კვდომის საფრთხისგან დაცვის მიზნების გათვალისწინებით ქართული და აფხაზური ენების ტექნოლოგიური ანბანების აგების ამოცანების სრულ და ამომწურავ გადაწყვეტებზე, რაც, მთლიანობაში, მიმართულია ერთიანი სახელმწიფოს ფარგლებში ქართველებისა და აფხაზების ევროკავშირში, უფრო ზოგადად კი, მომავლის კულტურულ სამყაროში ქართული და აფხაზური ენებით ღირსეული ანუ სრულფასოვანი შესვლის ეროვნულ-კულტურული მნიშვნელობების მქონე მიზნების უზრუნველყოფაზე.

მართლაც, სადისერტაციო კვლევით კ.ფხაკაძის ქართული ენის ლოგიკურ გრამატიკაზე დაყრდნობით წინა წლებში უკვე შემუშავებული ფხაკაძე-ჩიქვინიძისა და ფხაკაძე-ჩიჩუას მეთოდებზე და, ასევე, სხვადასხვა დღეს კარგად ცნობილი ინტრუმენტებსა და პლატფორმებზე დაყრდნობით, ერთი მხრივ, გაძლიერდა, მეორე მხრივ კი, შეიქმნა ქართული და აფხაზური ენების სრული ტექნოლოგიური უზრუნველყოფისთვის ანუ ქართული აფხაზური ენების სრულყოფილად მცოდნე ინტელექტუალური კომპიუტერული სისტემების შემუშავებისთვის, რაც იგივეა, ქართული და აფხაზური ენების ტექნოლოგიური ანბანის აგებისთვის აუცილებელი მთელი რიგი დამხმარე ინსტრუმენტები და ასეთი ინტრუმენტების ამგები მეთოდები და ამ გზით შემუშავდა ქართული უნივერსალური ჰკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია, რაც სადისერტაციო კვლევის მთავარი შედეგია. ანუ, სხვა სიტყვებით, ეს იმას ნიშნავს, რომ სადისერტაციო კვლევის ფარგლებში რიგ შემთხვევებში უკვე არსებული

სისტემების გაძლიერებით, რიგ შემთხვევებში კი არსებული სისტემებიდან სრულიად დამოუკიდებლად აიგო ქართული უნივერსალურ ჭკვიან კორპუსში ჩადგმული სახით ფუნქციონირებადი ისეთი მთელი რიგი ქართული და აფხაზური ენობრივი რესურსები და ტექნოლოგიები, როგორცაა, მაგალითად:

1. აფხაზური ენის თვითგანვითარებადი კორპუსი და მასში ჩადგმული აფხაზური ვებ-კლავიატურა;
2. აფხაზურ კორპუსში ჩადგმული პირველი და ჯერ-ჯერობით ერთადერთი აფხაზური მკითხველი სისტემა;
3. ჩეჩნური, ლეკური და ყაზარდოული ენების თვითგანვითარებადი კორპუსების პირველი საცდელი ვერსიები და მათში ჩადგმული ჩეჩნური, ლეკური და ყაზარდოული ვებ-კლავიატურები;
4. ქართული ტიტრირებული მეტყველების კორპუსის საცდელი ვერსია;
5. აფხაზური, ჩეჩნური, ლეკური და ყაზარდოული ტიტრირებული მეტყველების კორპუსების ამგები ინსტრუმენტები;
6. ქართულ უნივერსალურ ჭკვიან კორპუსში ჩადგმული მეგრული თვითგანვითარებადი კორპუსი;
7. მომხმარებლების მოწყობილობებიდან (კომპიუტერი, მობილური) კორპუსის გამაძლიერებელი ინსტრუმენტები;
8. ქართული და აფხაზური ტიტრირებული სამეტყველო მონაცემების სემენტატორების საცდელი სისტემები;
9. ქართული ჭკვიანი ფურცლის პირველი საცდელი ორენოვანი (ქართულ-აფხაზური) ვერსია (ორენოვანი ბეჭედა; ორენოვანი კითხვა; ორენოვანი ხმოვანი მართვა; ორენოვანი ორთოგრაფიული მართლწერა; ქართული სინაქსური მართლწერა) და მასში ჩადგმული სახით ფუნქციონირებადი:
 - 9.1. კორპუსში ჩადგმული ქართული თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსიის თითქმის სრული სამომხმარებლო ვერსია;
 - 9.2. კორპუსში ჩადგმული აფხაზური თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსია;
 - 9.3. კორპუსში ჩადგმული ქართული თვითგანვითარებადი გრამატიკული მართლწერის საცდელი ვერსიის საცდელ-სამომხმარებლო ვერსია;
10. ქართული ხმოვანმართვიანი მკითხველის გაფართოების გზით შემუშავებული ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები:
 - 10.1. თანამედროვე აფხაზური მეტყველების ამოცნობისა და ხმოვანი მართვის პირველი საცდელი ვერსიები;
 - 10.2. ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები;
11. ქართული N, A და V ტიპის სიტყვების ავტომატური მორფო-სინტაქსურ-სემანტიკური ანალიზატორის თითქმის სრული საცდელი და საცდელ-სამომხმარებლო ვერსიები;
12. ქართული ხმოვანი ლექსიკონის ახალი გაძლიერებული ვერსია;

13. ქართული ტექსტის ამოცნობის საცდელი სისტემები;
14. აფხაზური ტექსტის ამოცნობის პირველი საცდელი სისტემები;
15. ქართული სიტყვების სრული ავტომატური ანალიზატორის (ანუ ტაგერის, დესკრიპტორის, ლემატიზატორის, გენერატორისა და ანოტატორის) პირველი საცდელი ვერსია;
16. სადისერტაციო კვლევის მთავარი შედეგი ანუ ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია;

ყოველივე ეს ზემოაღნიშნული გასაგებს ხდის, რომ სადოქტორო კვლევა „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ თავისი შედეგებით მნიშვნელოვნად ზრდის ენობრივი რესურსებისა და ტექნოლოგიების თვალსაზრისით ქართული და აფხაზური ენების მხარდაჭერის მანამდე არსებულ მოცულობასა და ხარისხს. ეს კი, თავის მხრივ, გასაგებს ხდის, რომ სადისერტაციო კვლევის შედეგები ანუ ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია მეტად მნიშვნელოვანია ბუნებრივი ენების ტექნოლოგიების მიზნით მოქმედი სამეცნიერო სფეროსთვის (გამოთვლითი ლინგვისტიკა, ბუნებრივი ენების ლოგიკა, ხელოვნური ინტელექტი, საინფორმაციო და საკომუნიკაციო ტექნოლოგიები), კერძოდ კი, ქართული და აფხაზური ენების ტექნოლოგიების მიზნით მოქმედი სამეცნიერო სფეროებისათვის. ამასთან, სადისერტაციო კვლევის შედეგები უნდა შეფასდეს აგრეთვე ციფრული კვდომის იმ მაღალი ხარისხის საფრთხეების გათვალისწინებით, რომლის წინაშეცაა სხვა ამ მხრივ დაწინაურებულ ენებთან შედარებით მათემატიკურად არასაკმარისად შესწავლილი და, შესაბამისად, არასაკმარისად ტექნოლოგიზებული საქართველოს სახელმწიფო ენები - ქართული და აფხაზური.

Abstract

From technological support point of view, the Georgian and Abkhazian languages are alarmingly lag behind compared to any of those 21 European languages, which, according to the well-known research “Europe's Languages in the Digital Age” done by Meta-Net with financial support of Euro Structures are under the danger of digital extinction in the rapidly forthcoming digital age. - This clearly indicates the necessity of the overcoming this lagging as soon as it is possible.

Thus, PhD thesis “Elaboration of the New Developing Tools and Methods of the Georgian Smart Corpus and Improvement of Already Existing Ones”, which is a sub-project of a long-term projects “Technological Alphabet of the Georgian Language” and “Plan-Program for Complete technological support of Abkhazian Language” of the Scientific-Educational Center for Georgian Language Technology, aims at reducing this alarming lagging and, therefore, it is directly engaged with the aim of saving the Georgian and Abkhazian language from the danger of digital extinction and, also, with the aim of joining the European Union, in more general, in the future cultural world with the technologically completely supported Georgian and Abkhazian languages, in other words, with the Georgian and Abkhazian technological alphabets - with intellectual computer systems knowing Georgian and Abkhazian languages completely and perfectly.

Indeed, on the base of methods of Pkhakadze-Chikvinidze and Pkhakadze-Chichua, elaborated on the base the K.Pkhakadze’s Logical Grammar of the Georgian Language and, also, on the base of different today well-known tools and platforms within the PhD thesis are built such Georgian and Abkhazian language technology systems, which are necessary for creating Georgian and Abkhazian Technological Alphabets. They are:

1. Self-developing Corpus of Abkhaz language and Abkhazian web keyboard;
2. The first and only Abkhazian reader system inbuilt in the corpus;
3. The first trial versions of the self-developing corpuses of Chechen, Lezgian and Kabardian languages and web keyboards for Chechen, Lezgian and Kabardian languages;
4. Trial version of the Georgian titrated speech corpus;
5. Tools for the construction Abkhazian, Chechen, Lezgian and Kabardian titrated speech corpuses;
6. Mingrelian Self-Developing Corpus inbuilt in the Georgian Universal Corpus;
7. Tools for the strengthen of the corpus from the user’s devices (PC, mobile);
8. Trial segmentator systems for Georgian and Abkhaz titrated speech data;
9. The first bilingual (Georgian-Abkhazian) version of the Georgian Smart Paper (bilingual print, reading, voice management, spelling; Georgian syntax checker) and embedded in it:
 - 9.1. Almost complete applied version of the trial-applied version of the self developing Georgian orthographic spell-checker;

- 9.2. Trial-applied version of Abkhazian Self-Developing orthographic spell-checker;
- 9.3. Trial-applied version of the Georgian Self-Developing Grammar checker system;
- 10. Trial versions of old Georgian and modern Abkhazian voice managed reader systems developed by expanding Georgian voice managed reader system;
 - 10.1. First trial versions of Abkhaz speech recognizer and voice management systems
 - 10.2. Trial versions of old Georgian and modern Abkhazian voice managed reader systems;
- 11. Almost complete trial and trial-applied versions of the morpho-syntactic-semantic analyzer for Georgian N, A and V type words;
- 12. New strengthened Version of Georgian Voce Vocabulary;
- 13. Trial systems for Georgian text recognizer;
- 14. First trial systems for Abkhazian text recognizer;
- 15. The first trial version of the complete automated Georgian word analyzer (i.e tagers, descriptors, lematizers, generator and annotator);
- 16. The main result of the dissertation research, i.e., the first trial-applied version of the Georgian universal smart corpus.

All above-mentioned makes clear that from the language technology and resources points of view the results of PhD thesis “Elaboration of the New Developing Tools and Methods of the Georgian Smart Corpus and Improvement of Already Existing Ones” significantly increases the language support quality and capacity for Georgian and Abkhazian languages.

This makes clear that the results of the PhD thesis are very important for the scientific area (computational linguistics, artificial intelligence, information and communication technologies) which is engaged with natural language processing and, namely, for the scientific area which is engaged with technological processing of Georgian and Abkhazian languages.

At the same time, the importance of the results of this PhD thesis should be also estimated in the context of that high danger of the digital death which faces the Georgian and Abkhazian languages, which are insufficiently studied from mathematical points of view, and therefore, insufficiently technologically elaborated compared with other languages that are advanced ones from this points.

შინაარსი

შესავალი.....	17
თავი 1. ლიტერატურის მიმოხილვა.....	26
1.1. მოკლედ იმის შესახებ, თუ საიდან მომდინარეობს დისერტაციაში დასმული პრობლემა	26
1.2. მოკლედ იმის შესახებ თუ რატომ მივიჩნევთ ქართულ და აფხაზურ ენებს ციფრული კვდომის მაღალი საფეხურის საფრთხის ქვეშ მყოფ ენებად.....	30
1.3. კიდევ ერთხელ იმის შესახებ, თუ საიდან მომდინარეობს დისერტაციაში დასმული პრობლემა	36
1.4. იმ AR/122/4-105/14 პროექტის შესახებ, რომლის გამაგრძელებელი პროექტიცაა სადისერტაციო კვლევა	45
1.5. აღნიშვნათა თეორიისა და ქართული ენის ლოგიკური გრამატიკის ჩვენთვის სასაფუძვლო მნიშვნელობის საკითხები	50
თავი 2. შედეგები და მათი განსჯა.....	73
2.1. აფხაზური ენის თვითგანვითარებადი კორპუსი და მასში ჩადგმული აფხაზური ვებ კლავიატურა.....	73
2.2. აფხაზურ კორპუსში ჩადგმული პირველი და ჯერ-ჯერობით ერთადერთი აფხაზური მკითხველი სისტემის საცდელი ვერსია	77
2.3. ჩეჩნური, ლეკური და ყაზარდოული ენების თვითგანვითარებადი კორპუსების პირველი საცდელი ვერსიები და მათში ჩადგმული ჩეჩნური, ლეკური და ყაზარდოული ვებ-კლავიატურები.....	80
2.4. ქართული ტიტრირებული მეტყველების კორპუსის საცდელი ვერსია .	84
2.5. აფხაზური, ჩეჩნური, ლეკური და ყაზარდოული ტიტრირებული მეტყველების კორპუსების ამგები ინსტრუმენტები.....	88
2.6. ქართულ უნივერსალურ ჭკვიან კორპუსში ჩადგმული მეგრული თვითგანვითარებადი კორპუსი.....	92
2.7. მომხმარებლების მოწყობილობებიდან (კომპიუტერი, მობილური) კორპუსის გამაძლიერებელი ინსტრუმენტები.....	93
2.8. ქართული და აფხაზური ტიტრირებული სამეტყველო მონაცემების სეგმენტატორების საცდელი სისტემები	95
2.9. ქართული ჭკვიანი ფურცლის პირველი საცდელი ორენოვანი (ქართულ-აფხაზური) ვერსია (ორენოვანი ბეჭედა; ორენოვანი კითხვა; ორენოვანი ხმოვანი მართვა; ორენოვანი ორთოგრაფიული მართლწერა; ქართული სინტაქსური მართლწერა).....	98

2.9.1. კორპუსში ჩადგმული ქართული თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსიის თითქმის სრული სამომხმარებლო ვერსია	99
2.9.2. კორპუსში ჩადგმული აფხაზური თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსია.....	101
2.9.3. კორპუსში ჩადგმული ქართული თვითგანვითარებადი გრამატიკული მართლწერის საცდელი ვერსიის საცდელ-სამომხმარებლო ვერსია.....	102
2.10. ქართული ხმოვანმართვიანი მკითხველის გაფართოების გზით შემუშავებული ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები.....	109
2.10.1. თანამედროვე აფხაზური მეტყველების ამოცნობისა და ხმოვანი მართვის პირველი საცდელი ვერსიები.....	109
2.10.2. ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები.....	112
2.11. ქართული N, A და V ტიპის სიტყვების ავტომატური მორფო-სინტაქსურ-სემანტიკური ანალიზატორის თითქმის სრული საცდელი და საცდელ-სამომხმარებლო ვერსიები	118
2.12. ქართული ხმოვანი ლექსიკონის ახალი გაძლიერებული ვერსია.....	127
2.13. ქართული ტექსტის ამომცნობის საცდელი სისტემები	131
2.14. აფხაზური ტექსტის ამოცნობის პირველი საცდელი სისტემები	139
2.15. ქართული სიტყვების სრული ავტომატური ანალიზატორის (ანუ ტაგერის, დესკრიპტორის, ლემატიზატორის, გენერატორისა და ანოტატორის) პირველი საცდელი ვერსია.....	143
2.16. სადისერტაციო კვლევის მთავარი შედეგი ანუ ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია.....	151
დასკვნა	156
გამოყენებული ლიტერატურა.....	161

ცხრილების ნუსხა

ცხრილი 1 - ევროპის ენების ტექნოლოგიური მხარდაჭერების მახასიათებელი ცხრილი -----	36
ცხრილი 2 - ქართული და ევროპული ენების ტექნოლოგიური მხარდაჭერების შემადარებელი ცხრილი -----	38
ცხრილი 3 - ლოგიკური ბრუნების ცხრილი -----	62
ცხრილი 4 - ქართულ აფხაზური ტრანსლიტერაციის ცხრილი -----	77

სურათების ნუსხა

სურათი 1 - ენობრივი ყვავილი მათემატიკური გულით -----	58
სურათი 2 - ქართული ენა მის წერით, ზეპირ და სააზროვნო შრეებში -----	58
სურათი 3 - აფხაზური ვებ კლავიატურა -----	76
სურათი 4 - აფხაზური კორპუსი მოქმედებაში -----	76
სურათი 5 - აფხაზური კორპუსი ფრაზით ძიებისას -----	77
სურათი 6 - აფხაზური მკითხველი მოქმედებაში (1) -----	80
სურათი 7 - აფხაზური მკითხველი მოქმედებაში (2) -----	80
სურათი 8 - ჩეჩნური ვებ კლავიატურა -----	82
სურათი 9 - ლეკური ვებ კლავიატურა -----	83
სურათი 10 - ყაზარდოული ვებ კლავიატურა -----	83
სურათი 11 - სარეგისტრაციო გვერდი -----	85
სურათი 12 - ქართული ტიტრირებული კორპუსი -----	85
სურათი 13-14 - ტიტრირებული კორპუსში ხმის ჩაწერა -----	86
სურათი 15 - ხმის ჩამწერი ინსტრუმენტი -----	86
სურათი 16 - ჩაწერილი ხმები -----	87
სურათი 17 - ძებნა ტიტრირებულ მონაცემებში -----	88
სურათი 18 - ტიტრირებული მონაცემების გამოტანა -----	88
სურათი 19 - ტიტრირებულ კორპუსში ენების მხარდაჭერა -----	89
სურათი 20 - ჩეჩნური სიტყვის ძებნა -----	89
სურათი 21 - ტიტრირებული ჩანაწერის ჩვენება სიტყვაზე наной -----	89
სურათი 22 - ყაზარდოული სიტყვის ძებნა -----	90
სურათი 23 - ტიტრირებული ჩანაწერის ჩვენება სიტყვაზე Азмет -----	90
სურათი 24 - ლეკური სიტყვის ძებნა -----	91
სურათი 25 - ტიტრირებული ჩანაწერის ჩვენება სიტყვაზე Кассалий -----	91
სურათი 26 - ტიტრირებული აფხაზური ჩანაწერი სიტყვაზე Вари -----	91
სურათი 27 - ძებნა მეგრულ კორპუსში -----	92
სურათი 28 - ინფორმაცია სიტყვაზე „ჯიმა“ -----	93
სურათი 29 - გამაძლიერებელი ინსტრუმენტი კომპიუტერისთვის -----	94
სურათი 30-31-32 - გამაძლიერებელი ინსტრუმენტი მობლურისთვის -----	94
სურათი 33 - ქართული სეგმენტატორი კომპიუტერისთვის -----	95
სურათი 34 - აფხაზური სეგმენტატორი კომპიუტერისთვის -----	96
სურათი 35 - ქართული სეგმენტატორი მობილურისთვის -----	97
სურათი 36-37 - აფხაზური სეგმენტატორი მობილურისთვის -----	97
სურათი 38 - აფხაზური სეგმენტაციის მაგალითი -----	98
სურათი 39 - ქართული ორთოგრაფიული მართლმწერი (1) -----	99
სურათი 40 - ქართული ორთოგრაფიული მართლმწერი (2) -----	99
სურათი 41 - ქართული ორთოგრაფიული მართლმწერი (3) -----	100
სურათი 42 - ქართული ორთოგრაფიული მართლმწერის კორექტორი -----	100
სურათი 43 - აფხაზური ორთოგრაფიული მართლმწერი (1) -----	101
სურათი 44 - აფხაზური ორთოგრაფიული მართლმწერი (2) -----	101
სურათი 45 - აფხაზური ორთოგრაფიული მართლმწერი (3) -----	102

სურათი 46 - აფხაზური ორთოგრაფიული მართლმწერი (4) -----	102
სურათი 47 - სინტაქსური ანალიზის საწყისი გვერდი -----	103
სურათი 48 - სინტაქსური ანალიზი მოქმედებაში (1) -----	103
სურათი 49 - სინტაქსური ანალიზი მოქმედებაში (2) -----	104
სურათი 50 - სინტაქსური ანალიზი მოქმედებაში (3) -----	105
სურათი 51 - სინტაქსური ანალიზი მოქმედებაში (4) -----	105
სურათი 52 - სინტაქსური ანალიზი მოქმედებაში (5) -----	106
სურათი 53 - სინტაქსური ანალიზი მოქმედებაში (6) -----	106
სურათი 54 - სინტაქსური ანალიზი მოქმედებაში (7) -----	107
სურათი 55 - სინტაქსური ანალიზი მოქმედებაში (8) -----	107
სურათი 56 - სინტაქსური ანალიზი მოქმედებაში (9) -----	108
სურათი 57 - სინტაქსური ანალიზი მოქმედებაში (10) -----	108
სურათი 58 - ქართულ-აფხაზური ხმოვანი მართვა -----	109
სურათი 59 - ქართული ხმოვანი მართვის ბრძანებები -----	111
სურათი 60-61 - აფხაზური ხმოვანი მართვის ბრძანებები -----	111
სურათი 62 - ქართული კლავიატურა (კომპვერსია) -----	112
სურათი 63 - ქართული კლავიატურის კონფიგურაციის მოდელი -----	113
სურათი 64 - ძველი ქართული კლავიატურა (კომპ. ვერსია) -----	114
სურათი 65 - ძველი ქართული კლავიატურის კონფიგურაცია -----	114
სურათი 66 - აფხაზური კლავიატურის კომპ. ვერსია -----	114
სურათი 67 - აფხაზური კლავიატური კონფიგურაცია -----	115
სურათი 68 - ქართული მკითხველის კომპ. ვერსია -----	115
სურათი 69-70 - ქართული ხმოვანი ბრძანებების სტრუქტურა (1) -----	116
სურათი 71.1 - ქართული ხმოვანი ბრძანებების სტრუქტურა (2) -----	117
სურათი 71.2 - ძველი ქართული მკითხველის კომპ. ვერსია -----	117
სურათი 72 - აფხაზური მკითხველის კომპ. ვერსია -----	118
სურათი 73 - N ანალიზატორის გაძლიერებული ვერსია (1) -----	118
სურათი 74 - N ანალიზატორის გაძლიერებული ვერსია (2) -----	119
სურათი 75 - A ანალიზატორის გაძლიერებული ვერსია -----	119
სურათი 76 - V ანალიზატორის გაძლიერებული ვერსია (1) -----	120
სურათი 77 - V ანალიზატორის გაძლიერებული ვერსია (2) -----	121
სურათი 78 - V ანალიზატორის გაძლიერებული ვერსია (3) -----	121
სურათი 79 - V ანალიზატორის გაძლიერებული ვერსია (4) -----	122
სურათი 80 - V ანალიზატორის გაძლიერებული ვერსია (5) -----	122
სურათი 81 - V ანალიზატორის გაძლიერებული ვერსია (6) -----	123
სურათი 82 - V ანალიზატორის გაძლიერებული ვერსია (7) -----	124
სურათი 83 - N ტიპის ანოტატორის გაძლიერებული ვერსია (1) -----	125
სურათი 84-85 - N და A ტიპის ანოტატორების გაძლიერებული ვერსია ---	126
სურათი 86 - N ტიპის ანოტატორის გაძლიერებული ვერსია -----	126
სურათი 87 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (1) -----	127
სურათი 88 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (2) -----	128
სურათი 89 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (3) -----	128
სურათი 90 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (4) -----	129
სურათი 91 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (5) -----	130

სურათი 92 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (6) -----	130
სურათი 93 - ქართული ტექსტის გამომცნობი (1) -----	131
სურათი 94 - ქართული ტექსტის გამომცნობი (2) -----	131
სურათი 95 - ქართული ტექსტის ამომცნობი (3) -----	132
სურათი 96 - ამოსაცნობი ტექსტის სურათი -----	133
სურათი 97 - ქართული ტექსტის ამომცნობი (4) -----	133
სურათი 98-99 - ქართული ტექსტის ამომცნობი (5) -----	134
სურათი 100 - ქართული ტექსტის გამომცნობი (6) -----	135
სურათი 101-102 - ქართული ტექსტის ამომცნობი (7) -----	136
სურათი 103 - ქართული ტექსტის გამომცნობი (8) -----	137
სურათი 104 - ქართული ტექსტის გამომცნობი (9) -----	138
სურათი 105 - ქართული ტექსტის ამომცნობი (10) -----	138
სურათი 106-107 - ქართული ტექსტის ამომცნობი (11) -----	139
სურათი 108-109 - აფხაზური ტექსტის ამომცნობი (1) -----	140
სურათი 110 - აფხაზური ტექსტის ამომცნობი (2) -----	140
სურათი 111 - აფხაზური ტექსტის ამომცნობი (3) -----	141
სურათი 112 - აფხაზური ტექსტის ამომცნობი (4) -----	141
სურათი 113 - აფხაზური ტექსტის ამომცნობი (5) -----	142
სურათი 114 - აფხაზური ტექსტის ამომცნობი (6) -----	142
სურათი 115 - აფხაზური ტექსტის ამომცნობი (7) -----	142
სურათი 116 - აფხაზური ტექსტის ამომცნობი (8) -----	143
სურათი 117 - აფხაზური ტექსტის ამომცნობი (9) -----	143
სურათი 118 - კვანტორის ამომცნობი -----	143
სურათი 119 - ნაცვალსახელის ამომცნობი -----	144
სურათი 120 - კავშირების ამომცნობი -----	145
სურათი 121-122 - რიცხვითი სახელისა და შორისდებულის ამომცნობი --	145
სურათი 123-124 - ზმიზედისა და ნაწილაკის ამომცნობი -----	146
სურათი 125 - ზმნის ამომცნობი -----	147
სურათი 126-127 - კვანტორის და ნაცვალსახელის ამომცნობი -----	147
სურათი 128-129 - შორისდებულის და კავშირის ამომცნობი -----	148
სურათი 130-131 - ზედსართავის და არსებითის ამომცნობი -----	148
სურათი 132-133 - მრავალტიპიანი სიტყვების ამომცნობი -----	148
სურათი 134-135 - ზმნის და ნაწილაკის ამომცნობი -----	149
სურათი 136 - ზმნის ამომცნობი -----	149
სურათი 137 - ზმნის ამომცნობი (1) -----	150
სურათი 138 - ზმნის ამომცნობი (2) -----	150
სურათი 139 - ამოუცნობი სიტყვის ამომცნობი -----	150
სურათი 140 - სრული კორპუსის მონაცემები -----	155
სურათი 141 - ქართული კორპუსის მონაცემები -----	155
სურათი 142 - აფხაზური კორპუსის მონაცემები -----	155
სურათი 143 - ჩეჩნური კორპუსის მონაცემები -----	155
სურათი 144 - ყაზარდოული კორპუსის მონაცემები -----	155
სურათი 145 - ლეკური კორპუსის მონაცემები -----	155
სურათი 146 - მეგრული კორპუსის მონაცემები -----	155

მადლიერება

მადლობა მინდა გადავუხადო საქართველოს ტექნიკურ უნივერსიტეტის რექტორს აკადემიკოს არჩილ ფრანგიშვილს, რექტორის მოადგილეს პროფესორ ზურაბ გასიტაშვილსა და სოხუმის უნივერსიტეტის პროფესორს ბიძინა ლემონჯავას საქართველოს მეცნიერებათა აკადემიის მოამბეში ნაშრომის წარდგენისას გაწეული მხარდაჭერისთვის.

ასევე, მადლიერი ვარ შოთა რუსთაველის საქართველოს ეროვნულ სამეცნიერო ფონდის ჩემ PHDF-18-1228 სადოქტორო პროექტზე გაღებული ერთწლიანი საგრანტო მხარდაჭერისთვის, რის გარეშეც ცხადია ჩემი სადოქტორო კვლევა, ისეთი ნაყოფიერი, როგორც იყო, ვერ იქნებოდა.

გარდა ამისა, მადლობა მინდა ვუთხრა აგრეთვე ინფორმატიკისა და მართვის სისტემების ფაკულტეტის დეკანის მოადგილეს ქალბატონ თინა კაიშაურს, ჩემი სწავლების სამაგისტრო საფეხურიდან მოყოლებული უწყვეტი მხარდაჭერისათვის. ასევე, მადლობელი ვარ ხელოვნური ინტელექტის დეპარტამენტის ხელმძღვანელის მარიამ ჩხაიძისა და დეპარტამენტის სხვა ასევე ჩემთვის მეტად საპატივცემულო პროფესორებისა და ლექტორების, რომელთა საქმიანი შენიშვნები მე მეხმარებოდა შედეგების შემდგომ დახვეწაში.

მადლიერი ვარ ჩემი საგრანტო პროექტის საზღვარგარეთელი კონსულტანტების - დოქტორების ბაგჰერ ბაბაალის (შარიფის ტექნოლოგიების უნივერსიტეტის პროფესორი) ალექსანდრე მასხარაშვილის (გოტებორგის უნივერსიტეტის პოსტდოქტორანტ-მკვლევარი), ლაშა აბზიანიძის (გრონიგენის უნივერსიტეტის პოსტდოქტორანტ-მკვლევარი), რომლებთანაც გარდა ელექტრონული კონსულტაციებისა მე მქონდა ერთკვირიანი მჭიდრო საკონსულტაციო ურთიერთობები თბილისის პირველი საერთაშორისო საზაფხულო სკოლის „ლოგიკა, ენა, ხელოვნური ინტელექტი“ ფარგლებში. ასევე, მეტად მადლიერი ვარ ჩემი საგრანტო პროექტის ადგილობრივი კონსულტანტების - დოქტორების გიორგი ჩიჩუასი და მერაბ ჩიქვინიძის, რომელთა ხელშეწყობის გარეშე მე გამიჭირდებოდა კვლევით დასახული მიზნების შედეგიანი გასრულება. მადლიერი ვარ ასევე ჩემი ახალგაზრდა კოლეგის დავით კურცხალიასი, რომელთან ერთადაც მე მოწია მუშაობა შოთა რუსთაველის საქართველოს ეროვნული საეცნიერო ფონდის იმ წარმატებულ პროექტზე, რომლის გამაგრძელებელი პროექტიცაა ჩემი სადოქტორო კვლევა.

დაბოლოს, განსაკუთრებული მადლობა მინდა ვუთხრა ქართული ენის ტექნოლოგიების ცენტრის დირექტორს და ჩემი სადოქტორო კვლევისა და საგრანტო პროექტის სამეცნიერო ხელმძღვანელს პროფესორ კონსტანტინე ფხაკაძეს, რომლის გარეშეც არც ეს ნაშრომი შეიქმნებოდა და არც იმის საშუალება მომეცემოდა ჩავბმულიყავი ისეთ მნიშვნელოვან კვლევებში, რომლებიც მიმართულია ციფრულ ეპოქაში ქართული და აფხაზური ენების ციფრული კვდომის საფრთხისგან დაცვისა და ტექნოლოგიური განვითარების მიზნებზე.

შესავალი

დისერტაციაში დასმული პრობლემის მეცნიერული სიახლე და ამოსავალი საფუძვლები: სადოქტორო თემა „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ არის საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის გრძელვადიანი პროექტების „ქართული ენის ტექნოლოგიური ანბანი“ და „აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის გეგმა-პროგრამა“ კიდევ ერთი სასაფუძვლო მნიშვნელოვანი ქვეპროექტი და ის მიზნად ისახავდა უკვე არსებული ქართული ინტელექტუალური ვებ-კორპუსის ანუ ქართული ჭკვიანი კორპუსის გაძლიერების გზით ქართული უნივერსალური ჭკვიანი კორპუსის საცდელი ვერსიის აგებას [1-14, 16, 51].

ამასთან, ხაზგასასმელია, რომ კ.ფხაკაძის სამეცნიერო ხელმძღვანელობით გასრულებული ჩემი სადოქტორო თემა „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ და, ასევე, ამ სადოქტორო თემის შემადგენელი ჩემი PHDF-18-1228 სადოქტორო პროექტი „ქართული აფხაზური ენებით ევროკავშირში ანუ სადოქტორო თემა - ქართული უნივერსალური ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“, რომლის სამეცნიერო ხელმძღვანელი ასევე იყო კ.ფხაკაძე, არის გამაგრებელი პროექტი ზემოაღნიშნული გრძელვადიანი პროექტების სასაფუძვლოდ მნიშვნელოვანი AR/122/4-105/14 ქვეპროექტის „კიდევ ერთი ნაბიჯი მოსაუბრე ქართული თვითგანვითარებადი ინტელექტუალური კორპუსისაკენ“ [1], რომლის სამეცნიერო ხელმძღვანელი ასევე იყო კ.ფხაკაძე და რომელიც 2017 წელს შოთა რუსთაველის ეროვნულმა სამეცნიერო ფონდმა წლის წარმატებულ პროექტად გამოაცხადა.

საქმე ისაა, რომ ამ შოთა რუსთაველის ეროვნული სამეცნიერო ფონდისა და საქართველოს ტექნიკური უნივერსიტეტის ერთობლივი

დაფინანსებით განხორციელებული ამ ორწლიანი AR/122/4-105/14 პროექტის „კიდევ ერთი ნაბიჯი მოსაუბრე ქართული თვითგანვითარებადი ინტელექტუალური კორპუსისაკენ“ ფარგლებში ქართული და აფხაზური ენების ციფრული კვდომის საფრთხისგან დაცვის მიზნით შემუშავდა მოსაუბრე ქართული თვითგანვითარებადი ინტელექტუალური ვებ-კორპუსი ანუ, მოკლედ, ქართული ჭკვიანი კორპუსი და მასში ჩადგმული მთელი რიგი უნიკალური ანუ სრულიად ახალი ქართული და აფხაზური ენობრივი ტექნოლოგიები, რომელთა წყალობით ქართული და აფხაზური ენები ბევრად უფრო დაცულნი გახდნენ ციფრული კვდომის საფრთხისგან, ვიდრე ისინი მანამდე ანუ ამ პროექტის ამ შედეგებამდე იყვნენ [1].¹

ამგვარად, როგორც უკვე აღინიშნა, სადოქტორო თემა „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“, მისი სათაურიდან გამომდინარე, მიზნად ისახავდა ქართული ჭკვიანი კორპუსის გაძლიერებას მასში უკვე ჩადგმული სახით ფუნქციონირებადი ინსტრუმენტებისა და ტექნოლოგიების გაძლიერებისა და, ასევე, კორპუსში ჩადგმული სრულიად ახალი ინსტრუმენტებისა და ტექნოლოგიების შემუშავების გზით და, შესაბამისად, ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელი ვერსიის აგებას. აღნიშნულთან დაკავშირებით, დასაბუთებით უნდა ითქვას, რომ სადოქტორო თემით დასახული მიზნები მიღწეულია და ქართული ჭკვიანი კორპუსის ზემოაღნიშნული სახის გაძლიერების გზით დღეს უკვე აგებულია ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია, რომელიც გამომდინარე იქიდან, რომ არის ქართველური და კავკასიური ენების ტექნოლოგიურად ყველაზე უფრო მეტად მხარდაჭერილი მრავალენოვანი და მრავალმოდულური კორპუსი, ჩვენ ხელთ არსებული მონაცემებით, არის უნიკალური ანუ სხვა მსგავსი ანალოგების არ მქონე და აქედან გამომდინარე, თავისი არსით, ეს

¹ იხილეთ <https://rustaveli.org/ge/tsarmatebuli-proeqebi/kidev-erti-nabiji-mosaubre-kartuli-tvitganvitarebadi-intelektualuri-korpusisaken> მისამართზე განთავსებული მასალა.

ახლად შემუშავებული ქართული უნივერსალური ჭკვიანი კორპუსი არის მეტად მნივნელოვანი მეცნიერული მნიშვნელობის მქონე სიახლე განსაკუთრებით იმ სამეცნიერო წრეებისთვის, რომლებიც დაკავებულნი არიან ქართული და აფხაზური ენების სრული ტექნოლოგიური უზრუნველყოფის მიზნებით [1-14].

უფრო დეტალურად სადისერტაციო კვლევის ამოსავალი საფუძვლების, მეცნიერული სიახლისა და აქტუალობის შესახებ: საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრში, ცენტრის დირექტორის - კონსტანტინე ფხაკაძის ხელმძღვანელობით, რომელიც, როგორც უკვე აღინიშნა, ამავდროულად ჩემი სადოქტორო თემისა „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ და ჩემი PHDF-18-1228 სადოქტორო პროექტის „ქართული აფხაზური ენებით ევროკავშირში ანუ სადოქტორო თემა - ქართული უნივერსალური ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ სამეცნიერო ხელმძღვანელიცაა, მოქმედებს გრძელვადიანი პროექტები „ქართული ენის ტექნოლოგიური ანბანი“² (2012 წლიდან) და „აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის გეგმა-პროგრამა“ (2015 წლიდან), რომელთა მიზანია ქართული და აფხაზური მოაზროვნე, მოსაუბრე და მთარგმნელი მანქანების აგება, რაც უწინარესობით ცხადია ითხოვს ქართული და აფხაზური მოაზროვნე და მოსაუბრე მანქანების ანუ, რაც იგივეა, ქართული და აფხაზური ენების სრულყოფილად მცოდნე კომპიუტერული სისტემების, სხვა სიტყვებით, ქართული და აფხაზური ენის ტექნოლოგიური ანბანების აგებას [16-17]. ეს იმას ნიშნავს,

² ეს პროექტი შემუშავდა ივ.ჯავახიშვილის სახელობის თბილის სახელმწიფო უნივერსიტეტში კ.ფხაკაძის ხელმძღვანელობით მოქმედი სახელმწიფო-მიზნობრივი პროგრამით „კომპიუტერის სრული პროგრამულ-მომსახურეობითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ დასახული მიზნებისა და მიღწეული შედეგების საფუძველზე.

რომ ეს გრძელვადიანი პროექტები, დღეს, უკვე კარს მომდგარ და სანახევროდ უკვე ამოქმედებულ ციფრულ ეპოქაში, პირდაპირ უკავშირდება ქართული და აფხაზური ენების ციფრული კვდომის საფრთხისგან დაცვის საერთო ქართულ ეროვნულ მიზანსა და პასუხისმგებლობას [18-21]. უფრო მეტიც, პირდაპირ შეიძლება ითქვას, რომ ამ პროექტებით საბოლოოდ გადასაწყვეტად დასახული მიზნების ვერ მიღწევის შემთხვევაში ანუ, იმ შემთხვევაში, თუ ჩვენ ვერ შევძლებთ ქართული და აფხაზური ენების ტექნოლოგიური ანბანების შემუშავებასა და, შესაბამისად, ქართული და აფხაზური მოაზროვნე, მოსაუბრე და მთარგმნელი მანქანების აგებას, მაშინ ჩვენ ვერც ქართული და აფხაზური ენების დაცვას შევძლებთ ციფრული კვდომის დღეს მათ წინაშე მდგომი მეტად მაღალი საფეხურის საფრთხეებისგან. ამგვარად, ეს ყველაფერი, იმის გათვალისწინებით, სადოქტორო კვლევა „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ არის ამ ზემოთ მოკლედ მიმოხილული მეტად მაღალი ეროვნული მნიშვნელობების მქონე გრძელვადიანი პროექტების შემდგომი განვითარების მიმდინარე ეტაპისათვის ძირეულად მნიშვნელოვანი ქვეპროექტი, ცალსახად გასაგებს ხდის და უდაოდ ასაბუთებს აგრეთვე სადოქტორო კვლევის მეტად მაღალ ეროვნულ მნიშვნელობასა და აქტუალობას. გარდა ამისა, ზემოთ უკვე აღნიშნული გასაგებს ხდის იმასაც, რომ სადოქტორო კვლევა ძირეულად ეყრდნობა გრძელვადიანი პროექტების „ქართული ენის ტექნოლოგიური ანბანი“ და „აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის გეგმა-პროგრამა“ ფარგლებში წარმოებული წარმატებული კვლევებით უკვე მიღწეულ შედეგებს [1-16].

კერძოდ, ეს შედეგები მიღწეულია ამ გრძელვადიანი პროექტების ისეთი უკვე გასრულებული ქვეპროექტების ფარგლებში, როგორცაა:

1. შოთა რუსთაველის ფონდის ფუნდამენტური კვლევების პროექტი „ქართული ენის ლოგიკური გრამატიკის საფუძვლები და მისი გამოყენება საინფორმაციო ტექნოლოგიებში“ (ხელმძღვანელი კ.ფხაკაძე) [22-25].

2. საქართველოს ტექნიკური უნივერსიტეტის პროექტი „ქართული ენის ტექნოლოგიური ანბანის ასაგებად აუცილებელი რიგი სისტემების გაფართოებადი (სწავლებადი) საინტერნეტო ვერსიების შემუშავება“ (ხელმძღვანელი კ. ფხაკაძე).

3. შოთა რუსთაველის ფონდის სადოქტორო პროექტი „ქართული ენით ევროკავშირში ანუ სადოქტორო თემა - ქართული გრამატიკული მართლმწერი (ანალიზატორი)“ (დოქტორანტი მ.ჩიქვინიძე, ხელმძღვანელი კ.ფხაკაძე) [26].

4. შოთა რუსთაველის ფონდის სადოქტორო პროექტი „ქართული ენით ევროკავშირში ანუ სადოქტორო თემა - ქართული მეტყველების სინთეზი და ამოცნობა“ (დოქტორანტი გ.ჩიჩუა, ხელმძღვანელი კ.ფხაკაძე) [27].

5. შოთა რუსთაველის ფონდის პროექტი „კიდევ ერთი ნაბიჯი მოსაუბრე ქართული თვითგანვითარებადი ინტელექტუალური კორპუსისაკენ“ (ხელმძღვანელი კ. ფხაკაძე) [28-36].

როგორც უკვე აღინიშნა სადისერტაციო კვლევის მიზანი იყო თვითგანვითარებადი ქართული უნივერსალური ჭკვიანი ვებ-კორპუსის პირველი საცდელ-სამომხმარებლო ვერსიის შემუშავება უკვე არსებული ქართული და აფხაზური ენობრივი რესურსებისა და ტექნოლოგიების გაძლიერებისა და, ასევე, ქართულისა და აფხაზურისათვის სრულიად ახალი ენობრივი რესურსებისა და ტექნოლოგიების შექმნის გზით. კერძოდ, სადისერტაციო კვლევით კ.ფხაკაძის ქართული ენის ლოგიკურ გრამატიკაზე დაყრდნობით წინა წლებში უკვე შემუშავებული ფხაკაძე-ჩიქვინიძისა და ფხაკაძე-ჩიჩუას მეთოდებზე და, ასევე, სხვადასხვა დღეს კარგად ცნობილი ინტრუმენტებსა და პლატფორმებზე დაყრდნობით, ერთი მხრივ, გაძლიერდა, მეორე მხრივ კი, შეიქმნა ქართული და აფხაზური ენების სრული ტექნოლოგიური უზრუნველყოფისთვის ანუ ქართული აფხაზური ენების სრულყოფილად მცოდნე ინტელექტუალური კომპიუტერული სისტემების შემუშავებისთვის, რაც იგივეა, ქართული და აფხაზური ენების

ტექნოლოგიური ანბანის აგებისთვის აუცილებელი ქვემოჩამოთვლილი ენობრივი რესურსები და ტექნოლოგიები:

1. აფხაზური ენის თვითგანვითარებადი კორპუსი და მასში ჩადგმული აფხაზური ვებ-კლავიატურა;
2. აფხაზურ კორპუსში ჩადგმული პირველი და ჯერ-ჯერობით ერთადერთი აფხაზური მკითხველი სისტემა;
3. ჩეჩნური, ლეკური და ყაზარდოული ენების თვითგანვითარებადი კორპუსების პირველი საცდელი ვერსიები და მათში ჩადგმული ჩეჩნური, ლეკური და ყაზარდოული ვებ-კლავიატურები;
4. ქართული ტიტრირებული მეტყველების კორპუსის საცდელი ვერსია;
5. აფხაზური, ჩეჩნური, ლეკური და ყაზარდოული ტიტრირებული მეტყველების კორპუსების ამგები ინსტრუმენტები;
6. ქართულ უნივერსალურ ჭკვიან კორპუსში ჩადგმული მეგრული თვითგანვითარებადი კორპუსი;
7. მომხმარებლების მოწყობილობებიდან (კომპიუტერი, მობილური) კორპუსის გამაძლიერებელი ინსტრუმენტები;
8. ქართული და აფხაზური ტიტრირებული სამეტყველო მონაცემების სეგმენტატორების საცდელი სისტემები;
9. ქართული ჭკვიანი ფურცლის პირველი საცდელი ორენოვანი (ქართულ-აფხაზური) ვერსია (ორენოვანი ბეჭვდა; ორენოვანი კითხვა; ორენოვანი ხმოვანი მართვა; ორენოვანი ორთოგრაფიული მართლწერა; ქართული სინტაქსური მართლწერა) და მასში ჩადგმული სახით ფუნქციონირენადი:
 - 9.1. კორპუსში ჩადგმული ქართული თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსიის თითქმის სრული სამომხმარებლო ვერსია;
 - 9.2. კორპუსში ჩადგმული აფხაზური თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსია;

- 9.3. კორპუსში ჩადგმული ქართული თვითგანვითარებადი გრამატიკული მართლმწერის საცდელი ვერსიის საცდელ-სამომხმარებლო ვერსია;
10. ქართული ხმოვანმართვიანი მკითხველის გაფართოების გზით შემუშავებული ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები:
- 10.1. თანამედროვე აფხაზური მეტყველების ამოცნობისა და ხმოვანი მართვის პირველი საცდელი ვერსიები;
- 10.2. ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები;
11. ქართული N, A და V ტიპის სიტყვების ავტომატური მორფო-სინტაქსურ-სემანტიკური ანალიზატორის თითქმის სრული საცდელი და საცდელ-სამომხმარებლო ვერსიები;
12. ქართული ხმოვანი ლექსიკონის ახალი გამლიერებული ვერსია;
13. ქართული ტექსტის ამოცნობის საცდელი სისტემები;
14. აფხაზური ტექსტის ამოცნობის პირველი საცდელი სისტემები;
15. ქართული სიტყვების სრული ავტომატური ანალიზატორის (ანუ ტაგერის, დესკრიპტორის, ლემატიზატორის, გენერატორისა და ანოტატორის) პირველი საცდელი ვერსია;
16. სადისერტაციო კვლევის მთავარი შედეგი ანუ ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია;
- ყოველივე ეს ზემოაღნიშნული გასაგებს ხდის, რომ სადოქტორო კვლევა „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ თავისი შედეგებით მნიშვნელოვნად ზრდის ენობრივი რესურსებისა და ტექნოლოგიების თვალსაზრისით ქართული და აფხაზური ენების მხარდაჭერის მანამდე არსებულ მოცულობასა და ხარისხს, რაც, თავის მხრივ, კიდევ ერთხელ სრულიად ცალსახად ასაბუთებს როგორც სადისერტაციო თემით წარმოებული კვლევის მეცნიერული სიახლეს, ასევე ამ კვლევის მაღალ პროდუქტიულობას.

სადისერტაციო კვლევის პრაქტიკული და გამოყენებითი მნიშვნელობის მოკლე მიმოხილვა: უდაო, რომ მაღალტექნოლოგიური ბუნებრივი ენობრივი ტექნოლოგიების დღეს მეტად ინტენსიურად მიმდინარე კონსტრუირებისა და დამკვიდრების ყოველდღიური პროცესების ფონზე პრაქტიკული და გამოყენებითი თვალსაზრისებითაც ასევე მეტად აქტუალურია სადისერტაციო თემით წარმოებული კვლევის შედეგად საცდელი, სამომხმარებლო და საცდელ-სამომხმარებლო სახით უკვე შემუშავებული ქართული ენობრივი ტექნოლოგიური სისტემები (იხილეთ კვლევის შედეგების ზემოთ წარმოდგენილი ჩამონათვალი და ასევე სადისერტაციო ნაშრომის 2-ე თავი სათაურით - შედეგები და განსჯა). მართლაც, ამ უკვე შემუშავებული სისტემებითა და მათი შემდგომი მცირედ მოდიფიცირებული და გაუმჯობესებული ვერსიებით უკვე დღესვე სხვადასხვა მოცულობით პრაქტიკულად ისარგებლებს:

1. განათლების, მეცნიერების, კულტურისა და სპორტის სამინისტრო, რათა ახალ საფეხურზე აიყვანოს და, შესაბამისად, ხარისხობრივად გააუმჯობესოს: ა) საქართველოში არაქართულენობრივ, ხოლო თავად აფხაზეთის ავტონომიურ ოლქში არააფხაზურენოვან რეგიონებში ქართული და აფხაზური ენების სწავლისა და სწავლების პროცესი; ბ) დღეს ინტენსიურად ინტერნეტიზირებად საგანმანათლებლო სისტემაში თანამედროვე ტექნოლოგიების გამოყენებით ქართულის, აფხაზურისა და სხვა საგნების ქართულად და აფხაზურად სწავლების პროცესები; გ) ინტერნეტში მოზარდებისათვის ქართული და აფხაზური ენებისა და სასწავლო-საგანმანათლებლო პროგრამების ინტენსიური და მიზნობრივი მიწოდების ასევე მეტად მნიშვნელოვანი პროცესები.

2. დაბადებით ყრუ-მუნჯები ან ამა თუ იმ მიზეზით მეტყველების უნარ შეზღუდულები/დაკარგულები. - ამ თვალსაზრისით განსაკუთრებით საინტერესოა ქართული და აფხაზური ელ-ტექსტების მკითხველი ხმოვანმართვიანი აპლიკაციები, რომელთა დახმარებითაც შესაძლებელია ამ

პირების საზოგადოებასთან ინტეგრაციის ანუ გარემო პირებთან სასაუბრო ურთიერთობებში შესვლის პროცესების გაუმჯობესება.

3. ნებისმიერი დაწესებულება, რომელიც მოიხმარს ქართულ და აფხაზურ ენებს. - უკვე შემუშავებული ქართული და აფხაზური ტექნოლოგიური სისტემები დაეხმარება მათ იქ ქართული და აფხაზური ენებით მიმდინარე საქმეთა წარმოების ხარისხობრივ გაუმჯობესებაში.

4. ნებისმიერი კერძო პირი, ვინც მოიხმარს ან/და სწავლობს ქართულ და აფხაზურ ენებს;

თუმცა, ზემოთ უკვე აღნიშნულის გათვალისწინებით ხაზს ვუსვამთ, რომ სადისერტაციო თემით დაგეგმილი კვლევა იყო მიმართული არა უშუალო პრაქტიკულ შედეგებზე, მსგავსად ბიზნეს სტარტაპებისა, არამედ ის, ისე როგორც ეს ფუნდამენტური ხასიათის სადოქტორო თემების შემთხვევაშია, მიმართული იყო ინფორმატიკის, კერძოდ კი ლოგიკის, ენისა და ხელოვნური ინტელექტის ფარგლებში მდებარე ურთულესი და უმნიშვნელოვანესი ამოცანების გადაწყვეტაზე და აქედან გამომდინარე სადისერტაციო თემის ფარგლებში მიღწეული შედეგები გამოყენებითი თავლსაზრისებით მეტად მნიშვნელოვანია იმ სამეცნიერო დაწესებულებებისა და, ასევე, იმ პირებისათვის, რომლებიც კვლევით საქმიანობას ეწევიან ქართული და აფხაზური ენების სრული ტექნოლოგიური დამუშავების მიზნებითა და მიმართულებით.

თავი 1. ლიტერატურის მიმოხილვა

1.1. მოკლედ იმის შესახებ, თუ საიდან მომდინარეობს დისერტაციაში დასმული პრობლემა

საქართველოში, ქართული ენისა და მეტყველების დამუშავების მიზნით კვლევითი პროცესების გასული საუკუნის 60-70-იანი წლებიდან მიმდინარეობს [37-47]. თუმცა, ქართული ენის სრული ტექნოლოგიური უზრუნველყოფის ანუ ქართული ენის სრულყოფილად მცოდნე ინტელექტუალური კომპიუტერული სისტემის შემუშავების, რაც იგივე, ქართული ენის ტექნოლოგიური ანბანის აგების მიზნით კვლევების პირველი მიზნობრივი ეტაპი ფორმალურად დაიწყო 2000-2002 წლებში ივ.ჯავახიშვილის სახელობის თბილისის სახელმწიფო უნივერსიტეტის ი.ვეკუას სახელობის გამოყენებითი მათემატიკის ინსტიტუტში კ.ფხაკაძის ხელმძღვანელობით მოქმედი სახელმწიფო-მიზნობრივი პროგრამის „კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ ჯერ შემუშავებითა და, მერე უკვე, 2003 წლიდან, მისი ამოქმედებით. ქართული ენის სრული ტექნოლოგიური უზრუნველყოფისა და ციფრული კვდომის საფრთხისგან დაცვის მიზნით 2000-2003 წლებში მიმდინარე ეს მეტად მნიშვნელოვანი პროცესები განხორციელდა გამოყენებითი მათემატიკის ინსტიტუტის იმდროინდელი დირექტორის პროფესორ დავით გორდეზიანის, თბილისის სახელმწიფო უნივერსიტეტის იმდროინდელი რექტორის აკადემიკოს როინ მეტრეველის, საქართველოს პარლამენტის განათლების, მეცნიერების, კულტურისა და სპორტის კომიტეტის იმდროინდელი თავმჯდომარის აკადემიკოს ნოდარ ამაღლობელისა და საქართველოს იმდროინდელი პრეზიდენტის ედუარდ

შევარდნადის მხარდაჭერებითა და ქართული ეროვნული მნიშვნელობების მქონე მიზნებზე მიმართული მამულიშვილური განწყობებით [48-61].³

ქართული ენის სრული ტექნოლოგიური უზრუნველყოფის ანუ, როგორც უკვე აღინიშნა, ქართული ენის სრულყოფილად მცოდნე ინტელექტუალური კომპიუტერული სისტემის - ქართული ენის ტექნოლოგიური ანბანის აგების მიზნით კვლევების მეორე დღეს მიმდინარე ეტაპი იწყება 2011 წლის 8 თებერვალს საქართველოს ტექნიკურ უნივერსიტეტში უნივერსიტეტის რექტორის აკადემიკოს არჩილ ფრანგიშვილის ინიციატივითა და ქართული ენის ციფრული კვდომის საფრთხისგან დაცვის მიზნით კ.ფხაკაძის ხელმძღვანელობით მოქმედი „ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის“ ფორმირებით. ამასთან, ეს მეორე ეტაპი საბოლოო სახელდებული სახით გასრულდა 2012 წელს „ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრში“ ცენტრის დირექტორის - კ.ფხაკაძის ხელმძღვანელობით გრძელვადიანი პროექტის „ქართული ენის ტექნოლოგიური ანბანი“ ამოქმედებით, რომლისგანაც, 2015 წელს, ცალკე გრძელვადიანი პროექტის სახით გამოიყო „აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის გეგმა-პროგრამა“.

ამგვარად, შეიძლება ითქვას, რომ საქართველოს ტექნიკურ უნივერსიტეტში 2015 წლიდან იწყება საქართველოს სახელმწიფო ენების - ქართულისა და აფხაზურის სრული ტექნოლოგიური უზრუნველყოფის, სხვა სიტყვებით, ქართული და აფხაზური ენების სრულყოფილად მცოდნე ინტელექტუალური კომპიუტერული სისტემების შემუშავების, რაც იგივე, ქართული და აფხაზური ენების ტექნოლოგიური ანბანის აგების მიზნით ღიად მიმდინარე სახელდებული კვლევები, რასთანაც სრულ თანმთხვევაშია სადოქტორო თემის „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი

³ ქართული ენისადმი ამ იმდროინდელი განწყობების პირდაპირი დასტური იყო ციფრული ეპოქის პრობლემებისადმი 2002 წელს საქართველოს პარლამენტში ორგანიზებული ფართო საკონფერენციო ტიპის სადისკუსიო შეხვედრა.

ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ და PHDF-18-1228 სადოქტორო პროექტის „ქართული აფხაზური ენებით ევროკავშირში ანუ სადოქტორო თემა - ქართული უნივერსალური ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ კვლევით მიზნები. საქმე ისაა, რომ ჩემი სადოქტორო თემისა და ჩემი PHDF-18-1228 სადოქტორო პროექტის პირდაპირი მიზანი არის ქართული ინტელექტუალური ვებ-კორპუსის, სხვა სიტყვებით, ქართული ჭკვიანი კორპუსის გაძლიერების გზით ქართული უნივერსალური ჭკვიანი კორპუსის საცდელი ვერსიის აგება, რაც პირდაპირაა მიმართული ქართული და აფხაზური ენების ციფრული კვდომის საფრთხისგან დაცვის მიზნით ქართული და აფხაზური ენების სრულყოფილად მცოდნე ინტელექტუალური კომპიუტერული სისტემების შემუშავების, რაც იგივეა, ქართული და აფხაზური ენების ტექნოლოგიური ანბანის აგების ამოცანებზე.

ხაზგასასმელია ისიც, რომ 2019 წლის 17 მაისს საქართველოს მეცნიერებათა ეროვნული აკადემიის ინოვაციებისა და მაღალი ტექნოლოგიების ცენტრში საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრი, კერძოდ კი ცენტრის დირექტორი - პროფესორი ვ.ფხაკაძე წარსდგა მოხსენებით „ქართული და აფხაზური ენებით ევროკავშირში ანუ ქართული და აფხაზური ენების სრული ტექნოლოგიური უზრუნველყოფის მიზნები და პრობლემები“. მოხსენებისას, რომელიც ნაწილობრივ სადისერტაციო კვლევისთვის სასაფუძვლო მონოგრაფიული ნაშრომის „ქართული ინტელექტუალური კორპუსი: მიზნები, შედეგები, რეკომენდაციები“ [1] პრეზენტაცია იყო, ძირითად აქცენტირებული იყო ქართული და აფხაზური ენების ციფრული კვდომის საფრთხისგან დაცვის მიზნებსა და პრობლემებზე და ამის გათვალისწინებით გაკეთდა შემდეგი რეკომენდაციები: **პირველი:** საქართველოს სახელმწიფო ენების - ქართულისა და აფხაზურის ციფრული კვდომის საფრთხისგან დაცვის პასუხისმგებლობა ცალსახად გვკარნახობს

„საქართველოს სახელმწიფო ენების კულტურული დაცვისა და ტექნოლოგიური განვითარების სამეცნიერო-კვლევითი ინსტიტუტის“ ფორმირების აუცილებლობას; **მეორე:** შოთა რუსთაველის საქართველოს ეროვნულმა სამეცნიერო ფონდმა „სახელმწიფო ენების კულტურული დაცვა და ტექნოლოგიური განვითარება“ უნდა განსაზღვროს ცალკე მდგომ უმაღლესი საფეხურის ინტერდისციპლინურ პრიორიტეტულ მიმართულებად; **მესამე:** აღნიშნულთა გათვალისწინებით წელიწადში ორჯერ (14 აპრილს - ქართული ენის დღისადმი მიძღვნილ და, ასევე, 27 ოქტომბერს - აფხაზური ენის დღისადმი მიძღვნილ) უნდა დაორგანიზდეს კონფერენცია პირობითი დასახელებით „საქართველოს ენები ციფრული ეპოქაში: კულტურული და ტექნოლოგიური ასპექტები“.⁴ - ამ ყველაფერთან დაკავშირებით, ხაზს ვუსვამთ იმას, რომ სამივე ეს ზემოწარმოდგენილი სარეკომენდაციო მოსაზრება სრულად იქნა მხარდაჭერილი საქართველოს მეცნიერებათა ეროვნული აკადემიის ინოვაციებისა და მაღალი ტექნოლოგიების ცენტრის 2019 წლის 17 მაისის #53 ოქმით, რადგან მოხსენებისას დასაბუთდა ისიც, რომ ამ რეკომენდაციების გათვალისწინება გასაგებს გახდის, რომ ამ დღეს უკვე კარს მომდგარ და სანახევროდ უკვე ამოქმედებულ ციფრულ ეპოქაში ქართული სახელმწიფო ზრუნავს ქართული და აფხაზური ენების დაცვისა და შენარჩუნებისათვის, რაც ოკუპირებულ ტერიტორიებზე მცხოვრებ აფხაზ ძმებთან მნიშვნელოვანი დიალოგის წამოწყებისა და მისი შემდგომი გაძლიერების ურყევ ნიადაგს მოგვცემს, რის გარეშეც ცხადია შუძმლებელი იქნება ქართულ-აფხაზური ურთიერთობების დღეს სამწუხაროდ ჩატეხილი ხიდის აღდგენა.

ამგვარად, ზემოაღნიშნულთა გათვალისწინებით, ხაზს ვუსვამთ იმას, რომ ჩემი სადოქტორო თემით დასმული პრობლემა მომდინარეობს ციფრული კვდომის იმ მაღალი საფეხურის საფრთხეებიდან, რომელთა წინაშეა დღეს საქართველოს სახელმწიფო ენები - ქართული და აფხაზური.

⁴ აქ დასახელებები პირობითია, თუმცა თითქმის სრულად ასახავს იმ მიზნებს, რასაც წარმოდგენილი რეკომენდაციები თავისი შინაარსებით ითვალისწინებს.

1.2. მოკლედ იმის შესახებ თუ რატომ მივიჩნევთ ქართულ და აფხაზურ ენებს ციფრული კვდომის მაღალი საფეხურის საფრთხის ქვეშ მყოფ ენებად

ქართული და აფხაზური ენების ბედთან დაკავშირებით ეს ზემოთ გამოთქმული საგანგაშო განცხადება და მისი ქვემოთ წარმოდგენილი დამასაბუთებელი ხედვები ეყრდნობა იმ მონაცემებს, რაც ქართული და აფხაზური ენებისთვის ჩვენ ხელთ გვაქვს [62–65] პუბლიკაციებით, ხოლო ევროპული ენებისთვის პუბლიკაციებით. საქმე ისაა, რომ მეტა-ქსელმა (<http://www.meta-net.eu/>) 2012 წლის 20 სექტემბერს, ენების ევროპული დღისადმი მიძღვნილ, ევროკავშირის 28-ვე სახელმწიფო ენაზე გამოაქვეყნა პრეს-რელიზი „სულ ცოტა 21 ევროპული ენაა ციფრული კვდომის საფრთხის ქვეშ“. პრესრელიზის მიზანი იყო ევროპული საზოგადოების ინფორმირება ევროპის ენების ციფრული კვდომის მაღალი საფეხურის საფრთხის თაობაზე. კერძოდ, პრესრელიზის თანახმად: *„ევროპამ უნდა მიიღოს ზომები, რათა მოამზადოს თავისი ენები ციფრული ეპოქისათვის. ისინი ძალიან მნიშვნელოვანი შემადგენლებია ჩვენი კულტურული მემკვიდრეობისა და, შესაბამისად, ცხადია, იმსახურებენ სამომავლო დაცულობას. ენების ევროპული დღე - 26 სექტემბერი კიდევ ერთხელ ადასტურებს ჩვენი კონტინენტის მდიდარი ენობრივი და კულტურული მემკვიდრეობის დაცვისა და განვითარების მეტად მაღალ მნიშვნელობას. ამასთან, მეტა-ქსელის კვლევა გასაგებს ხდის, როგორც იმას, რომ ციფრულ ეპოქაში, მრავალენოვანი ევროპა და მისი ლინგვისტური მემკვიდრეობა მნიშვნელოვანი გამოწვევების წინაშეა, ასევე იმასაც, რომ, დღეს, თანამედროვე ევროპის წინაშეა სრულიად ახალი გზები და სრულიად ახალი შესაძლებლობები. 200-ზე მეტი ექსპერტის მიერ მომზადებული და მეტა-ქსელის თეთრი ფურცლების სერიის 30 ტომად გამოცემული კვლევა, აფასებს კვლევაში მონაწილე თითოეული ენის ტექნოლოგიურ მხარდაჭერას ოთხი მიმართულებით. ესენია: ავტომატური თარგმანი, სასაუბრო ურთიერთობები, ტექსტის ანალიზი და ენობრივი რესურსები. ექსპერტების მიერ 30-დან 21 ენის (ანუ 70%-ის) მხარდაჭერა ზემოთ დასახელებული ოთხი*

მიმართულებიდან რომელიღაცა ერთში მაინც შეფასდა უდაბლესი კატეგორიით ანუ როგორც „სუსტი ან არანაირი მხარდაჭერა“. რამდენიმე ენა, ასეთებია ისლანდიური, ლატვიური, ლიტვური, და მალტური, ამ ოთხივე მიმართულებით უდაბლესი კატეგორიით შეფასდა. ამავდროულად, არცერთი ენა არ შეფასდა უმაღლესი კატეგორიით ანუ როგორც „სრულყოფილად მხარდაჭერილი“. მხოლოდ ინგლისური ენა შეფასდა როგორც „კარგად მხარდაჭერილი“, რომლის მერე მოდიან ისეთი „ზომიერად მხარდაჭერილი“ ენები, როგორიცაა პოლანდიური, ფრანგული, გერმანული, იტალიური და ესპანური. მხარდაჭერა ენებისა, როგორცაა ბასკური, ბულგარული, კატალონიური, ბერძნული, უნგრული და პოლონური, შეფასდა როგორც „ფრაგმენტული მხარდაჭერა“, რაც ამ ენებსაც ადგილს უჩენს მაღალი საფრთხის ქვეშ მყოფ ენათა ჯგუფში.“ [13] - ამგვარად, მათი თქმით, ის ენა, რომელსაც შესადარებელი ოთხი მიმართულებიდან (ეს მიმართულება - ავტომატური თარგმანი, სასაუბრო ურთიერთობები, ტექსტის ანალიზი და ენობრივი რესურსები) ერთში მაინც აქვს უდაბლესი ანუ „სუსტი ან არანაირი მხარდაჭერა“, ციფრული გაქრობის ანუ კვდომის საფრთხის ქვეშაა. გარდა ამისა, პრეს-რელიზით ბასკური, ბულგარული, კატალონიური, ბერძნული, უნგრული და პოლონური ენები, მიუხედავად იმისა, რომ მათ „ფრაგმენტული მხარდაჭერა“ აქვთ, ციფრული კვდომის საფრთხის ქვეშ მყოფ ენათა ჯგუფში ექცევიან. ანუ, ევროკომისიის დაფინანსებითა და 200-ზე მეტი ექსპერტის მონაწილეობით ჩატარებული კვლევის შედეგებზე დაყრდნობით თავად ევროპელები მიიჩნევენ, რომ კვლევაში მოქცეული 31 ევროპული ენიდან ციფრული კვდომის საფრთხის ქვეშ მყოფი ენებია „ფრაგმენტული“ და/ან „სუსტი ან არანაირი“ მხარდაჭერის მქონე ბასკური, ბულგარული, კატალონიური, ხორვატული, ჩეხური, დანიური, ესტონური, ფინური, გალიციური, ბერძნული, უნგრული, ისლანდიური, ირლანდიური, ლატვიური, ლიტვური, მალტური, ნორვეგიული, პოლონური, პორტუგალიური, რუმინული, სერბიული, სლოვაკური, სლოვენური, შვედური და უელსური ენები [13].

ამგვარად, იმის გათვალისწინებით, რომ ტექნოლოგიური მხარდაჭერის თვალსაზრისით ქართული, მით უფრო აფხაზური ენა ამ ზემოჩამოთვლილ საფრთხის ქვეშ მყოფ ევროპულ ენებთან შედარებით კიდევ უკანაა, გასაგებს ხდის, რომ საგანგაშო ზარი, რაც ევროპის ენებთან მიმართებაში, 2012 წელს გაისმა, დღეს, 2019 წელს, ქართულ და აფხაზურ ენებთან მიმართებაში, კიდევ უფრო ხმამაღლა უნდა გავაჟღეროთ, რადგან, ცხადია, რომ მხოლოდ ქართული სახელმწიფოსა და საზოგადოების ერთობლივი ძალისხმევით და კარგად დაგეგმილი ღონისძიებების გატარების შემთხვევაშია შესაძლებელი ქართულ და აფხაზურ ენებს ავაშოროთ დღეს მათ წინაშე მდგარი ციფრული კვდომის ეს განსაკუთრებით მაღალი საფრთხის საფრთხეები.

იმ მიზნით, რომ ქართულ და აფხაზურ ენასთან დაკავშირებით ზემოთ უკვე წარმოდგენილი საგანგაშო განწყობა კიდევ უფრო საგრძნობი და სარწმუნო გახდეს, ქვემოთ მოგვყავს ამონარიდები ნაშრომიდან „სტრატეგიული კვლევითი გეგმა 2020 წლის მრავალენოვანი ევროპისათვის“, რომელიც მეტა-ქსელის ტექნოლოგიურმა საბჭომ - 2012 წლის 1 დეკემბერს გამოაქვეყნა.

ევროკომისია: *„მრავალი წლის განმავლობაში და ათობით ენაში ვცდილობდით რა მედიის ავტომატურად მანალიზებული სისტემების შემუშავებას, ჩვენ მტკივნეულად განვიცდიდით ტექსტის მანალიზებული ინსტრუმენტებისა და რესურსების ნაკლებობას უმეტესი ენებისათვის. მეტა-ქსელის ანალიზი ძალიან ზუსტია. ენობრივი ტექნოლოგიები მართლაც სასაფუძვლო შემადგენელია ახალი თაობის საინფორმაციო ტექნოლოგიებისათვის. ენა, რომლისთვისაც არ იარსებებს ეს სასაფუძვლო ენობრივი ინსტრუმენტები და რესურსები უახლოეს მომავალში ვეღარ მიიღებს მონაწილეობას უკვე დაგეგმილ მთავარ ტექნოლოგიურ განვითარებებში.“* - რალფ შტაინბერგერი (საერთო ევროპული კვლევითი ცენტრი) [13].

ლატვია: „ისეთი მცირე ენებისათვის, როგორცაა ლატვიური, სასიცოცხლოდ მნიშვნელოვანია მხარი აუბას მზარდი ტემპებით მიმდინარე ტექნოლოგიურ პროგრესს. ჩვენი ენის სამომავლო არსებობის უზრუნველყოფის ერთადერთი გზა არის მისი მომხმარებლების უზრუნველყოფა ისეთივე შესაძლებლობებით, როგორი შესაძლებლობებიც აქვთ უფრო მრავალრიცხოვანი ენების მომხმარებლებს. ამიტომ თანამედროვე ტექნოლოგიების მოწინავე პოზიციებზე ყოფნა ჩვენი მიზნების მისაღწევი ერთადერთი გზაა.“ - ვალდის დომბროვსკისი (ლატვიის პრემიერ მინისტრი) [13].

დანია: „თუ ჩვენი მიზანია ვისარგებლოთ დანიური ენით მომავლის ტექნოლოგიურ სამყაროში, ახლავე უნდა მივიღოთ ზომები ჩვენ ხელთ არსებული ცოდნისა და გამოცდილების გაერთიანებისა და შემდგომი განვითარებისათვის. - ეს განსაკუთრებით ცხადი ხდება მეტა-ქსელის მოხსენების გათვალისწინებით. წინააღმდეგ შემთხვევაში ჩვენ ვდგებით იმ საფრთხის წინაშე, რომ მხოლოდ ინგლისურის სრულყოფილად მცოდნე ადამიანები ისარგებლებენ ახალი თაობის საინტერნეტო, სამობილურო და სარობოტო ტექნოლოგიებით, რომლებიც უკვე კარს მოგვადგა.“ - კირხმაიერ ანდერსენი (დანიური ენის საბჭოს დირექტორი) [13].

პორტუგალია: „ენობრივი ტექნოლოგიების სფეროში პორტუგალიური ენისთვის წარმოებული კვლევები უაღრესად მნიშვნელოვანია საინფორმაციო სამყაროში პორტუგალიურის გლობალური კომუნიკაციის ენად ფორმირებისათვის.“ - დოქტორი პედრო პასოს კოელო (პორტუგალიის პრემიერ მინისტრი) [13].

ჩეხეთის რესპუბლიკა: „მეტა-ქსელს შეაქვს მნიშვნელოვანი წვლილი ევროპის ენების ტექნოლოგიურ მხარდაჭერაში და, ასევე, მას ექნება შეუცვლელი როლი მრავალენოვანი ევროპული კულტურისა და საზოგადოების ფორმირებაში.“ - ივანე ვილჰელმი (განათლების, ახალგაზრდობისა და სპორტის მინისტრის მოადგილე) [13].

პოლონეთი: „ენობრივ ტექნოლოგიებს ... ექნება უზარმაზარი გავლენა თანამედროვე მსოფლიოს შესაძლებლობებსა და საკომუნიკაციო მოდელებზე, ისევე როგორც ისეთ ადამიანურ - ბუნებრივ ენებზე, როგორცაა პოლონური, რომელიც მონაწილეობას იღებს ამ პროცესში.“ - მიხეილ კლაიბერი (პოლონეთის მეცნიერებათა აკადემიის პრეზიდენტი)[13].

საბერძნეთი: „ენობრივი ტექნოლოგიების შემდგომი მხარდაჭერა ბერძნულ ენას და კულტურას არსებობას შეუნარჩუნებს ციფრულ სამყაროში ...“ - გიორგი ბაზბინიოტისი (განათლების, გახანგრძლივებული სწავლებისა და რელიგიურ საქმეთა მინისტრი) [13].

მალტა: „ ... მალტური ენის ტექნოლოგიური მხარდაჭერა შესაძლებლობას მისცემს ჩვენ ენას დაიკავოს ისეთივე მოწინავე ადგილი, როგორც აქვთ სხვა ენებს და იყოს მუდმივ განვითარებასა და გამოყენებაში.“ - დოლორეს კრისტინა (განათლებისა და დასაქმების მინისტრი) [13].

ლიტვა: „მომავალი თაობებისათვის ლიტვური ენის შენარჩუნება საერთო ევროპული პასუხისმგებლობაა. ლიტვური ენის მომავალი დიდადაა დამოკიდებული იმაზე, შევძლებთ თუ ვერა მივყვეთ საინფორმაციო ტექნოლოგიების შემდგომი განვითარების პროცესებს.“ - ანდრიუს კიბილიუსი (ლიტვის რესპუბლიკის პრემიერ მინისტრი) [13].

ირლანდია: „ენობრივი ტექნოლოგიები უკვე აღარ არის უმეტესი ევროპული ენების ფუფუნების საგანი - ეს უკვე არის ნებისმიერ შესაძლო სფეროში, დაწყებული ბიზნესიდან დამთავრებული ხელოვნებით, ამ ენების სიცოცხლისუნარიანობის გადარჩენისა და შენარჩუნების ერთადერთი საფუძველი, და ეს ასეა როგორც ირლანდიურისათვის, ასევე ნებისმიერი სხვა ევროპული ენისათვის.“ - ფერდი მაკ ფაილაი (ირლანდიის ფონდის აღმასრულებელი დირექტორი) [13].

უნგრეთი: „მეტა-ქსელს მნიშვნელოვანი წვლილი შეაქვს ევროპულ ინოვაციებსა, კვლევებსა, განვითარებებსა და ევროპული იდეის ეფექტურ განხორციელებაში.“ - ვალერია ჩიპე (უნგრეთის მეცნიერებათა აკადემიის გენერალური მდივნის მოადგილე) [13].

შვედეთი: „მაღალი ხარისხის ენობრივ ტექნოლოგიებს შეუძლიათ იქცნენ ევროპაში ენობრივი მრავალფეროვნების შენარჩუნების ყველაზე ეფექტურ საშუალებებად. მათ შეუძლიათ ყველა ენა სრულყოფილად გამოყენებადი გახადონ თანამედროვე საზოგადოებაში, რაც, თავის მხრივ, დემოკრატიის საკითხიცაა. ამ მიმართულებით მეტა-ქსელი თავის თავზე იღებს ცენტრალურ, უფრო მეტიც - გადამწყვეტ დატვირთვას.“ - ლენა ეკბერგი (შვედური ენის საბჭოს ხელმძღვანელი) [13].

სლოვენია: „უკიდურესად აუცილებელია სლოვენური ენისათვის ენობრივი ტექნოლოგიების უწყვეტი განვითარება, თუ რა თქმა უნდა გვინდა, რომ სლოვენურმა მომავლის ციფრულ სამყაროშიც შეინარჩუნოს თავისი განუწყვეტელი განვითარების გზა.“ - დოქტორი დანილო ტორკი (სლოვენიის რესპუბლიკის პრეზიდენტი) [13].

დიდი ბრიტანეთი: „მეტა-ქსელის საქმიანობა არის მნიშვნელოვანი ნაბიჯი იმ მომავლისაკენ, რომელშიც ენობრივი ტექნოლოგიები იქნება ყველა ჩვენგანის გარშემო, და რომლებიც, მიუხედავად იმისა ვსაუბრობთ თუ არა ჩვენ ერთსა და იმავე ენაზე, თანამშრომლობის, ბიზნესის წარმოებისა და მეგობრებსა და კოლეგებთან ცოდნის გაზიარების თავისუფალ საშუალებებს მოგვცემს.“ - დავით უილეტსი (უნივერსიტეტებისა და მეცნიერების სახელმწიფო მინისტრი) [13].

ესტონეთი: „თუ ჩვენ არ განვახორციელებთ ენობრივი ტექნოლოგიების განვითარების გეგმას და თუ არ ვითანამშრომლებთ ამავე მიმართულებით მოძრავ სხვა ენებთან, სამომავლოდ, ესტონური გაირიყება მომავლის საინფორმაციო სამყაროდან.“ - ესტონური ენის განვითარების 7 წლიანი გეგმა (2011-2017) [13].

ამგვარად, ეს ყველაფერი დასაბუთებულად გასაგებს ხდის რომ საქართველოს სახელმწიფო ენები - ქართული და აფხაზური მართლაც არიან ციფრული კვდომის მეტად მაღალი საფეხურის საფრთხის ქვეშ და მართლაც ითხოვენ ქართული სახელმწიფოს მხრიდან შესაბამის ყურადღებას. ეს კი, შესაბამისად, გასაგებს ხდის აგრეთვე „საქართველოს სახელმწიფო ენების

კულტურული დაცვისა და ტექნოლოგიური განვითარების სამეცნიერო-კვლევითი ინსტიტუტის“ ფორმირების გადაუდებელ აუცილებლობას, რითაც ქვეყანაში საქართველოს სახელმწიფო ენების - ქართული და აფხაზური ენების ტექნოლოგიური ანბანის აგების მიზნით 2015 წლიდან სახელდებით მიმდინარე კვლევების მეორე ეტაპი გადავა მესამე - დამაგვირგვინებელ ეტაპზე. - ჩვენ ამ კვლევების ამ მესამე ეტაპს დამაგვირგვინებელ ეტაპს ვუწოდებთ, რადგან, როცა ქვეყნის სახელმწიფო ენები ტექნოლოგიურად არასაკმარისი მხარდაჭერის გამო ციფრული კვდომის ანუ, რაც იგივეა, კულტურული ფუნქციების დაკარგვის მაღალი საფეხურის საფრთხის ქვეშ არიან, ზემოაღნიშნული სახის სამეცნიერო კვლევითი ინსტიტუტის ფორმირების გადაუდებელი აუცილებლობა სრულიად ნათელია!

1.3. კიდევ ერთხელ იმის შესახებ, თუ საიდან მომდინარეობს დისერტაციაში დასმული პრობლემა

ქვეთავში, წარმოდგენილია ქართული, აფხაზური და ევროპული ენების ტექნოლოგიური მხარდაჭერების შედარებითი ანალიზი და ამ ანალიზის საგანგაშო შედეგები, იმ მიზნით, რომ კიდევ უფრო გასაგები გავხადოთ ქართული და აფხაზური ენების წინაშე არსებული საფრთხეების რეალობა.

ამ მიზნით ქვემოთ მოგვყავს ცხრილი_1-ის გასააზრებლად აუცილებელი განმარტებები. კერძოდ, ცხრილი_1-ის სვეტში დასახელებით:

1. „სრულყოფილი“ თავმოყრილია სრულყოფილად მხარდაჭერილი ენები;
2. „კარგი“ თავმოყრილია კარგად მხარდაჭერილი ენები;
3. „საშუალო“ თავმოყრილია საშუალოდ მხარდაჭერილი ენები;
4. „ფრაგმენტული“ თავმოყრილია ფრაგმენტულად მხარდაჭერილი ენები;

5. „სუსტი/არანაირი“ თავმოყრილია სუსტად ან არანაირად მხარდაჭერილი ენები.

ასევე, ევროპული ენების ეს სხვადასხვა ხარისხის მხარდაჭერები გაიზომა ენობრივი ტექნოლოგიების ისეთი ოთხი სასაფუძვლოდ მნიშვნელოვანი მიმართულებით, როგორცაა [13]:

1. მანქანური თარგმანი - ამ მიმართულებით გათვალისწინებულია მთარგმნელი ტექნოლოგიების ხარისხი, ხელმისაწვდომობა და თარგმნით მიმართებაში ჩაყენებულ ენათა რაოდენობა; ენათაშორისი ანუ პარალელური კორპუსების ხარისხი, მოცულობა, მრავალფეროვნება და ხელმისაწვდომობა;

2. მეტყველების დამუშავება/სამეტყველო ურთიერთობები - ამ მიმართულებით გათვალისწინებულია მეტყველების ამომცნობი და მასინთეზებელი ტექნოლოგიების ხარისხი; სხვადასხვა სახის სამეტყველო კორპუსების რიცხვი და მოცულობა; სასაუბრო ანუ სამეტყველო ურთიერთობებზე დამყარებული ტექნოლოგიების რაოდენობა, მრავალფეროვნება და ხელმისაწვდომობა;

3. ტექსტის ანალიზი - ამ მიმართულებით გათვალისწინებულია ტექსტების მორფოლოგიურად, სინტაქსურად და სემანტიკურად მანალიზებელი ტექნოლოგიების ხარისხი, ხელმისაწვდომობა, რაოდენობა და მრავალფეროვნება; ანოტირებული კორპუსების მოცულობა და ხარისხი; უკვე არსებული ლექსიკური წყაროებისა და გრამატიკების ხარისხი და სისრულე;

4. რესურსები - ამ მიმართულებით გათვალისწინებულია უკვე არსებული ტექსტების კორპუსების, სამეტყველო კორპუსებისა და პარალელური კორპუსების ხარისხი და მოცულობა; ლექსიკური წყაროებისა და გრამატიკების ხარისხი და სისრულე.

ამგვარად, ცხრილ_1-ში ტერმინებით „მანქანური თარგმანი ანუ ავტომატური თარგმანი“, „მეტყველების დამუშავება ანუ სამეტყველო ურთიერთობა“, „ტექსტის ანალიზი“, „ენობრივი რესურსები“

დასახელებული სტრიქონები შინაარსობრივად ამ ტერმინების ზემოთ მოცემული განმარტებების შესაბამისად უნდა იქნეს გაგებული [13].

მხარდა-ჭერის ხარისხი	სრული	კარგი	საშუალო	ფრაგმენტული	სუსტი/არანაირი
მანქანური თარგმანი ანუ ავტომატური თარგმანი		ინგლისური	ფრანგული, ესპანური	კატალონიური, ჰოლანდიური, გერმანული, უნგრული, იტალიური, პოლონური, რუმინული	ბასკური, ბულგარული, ხორვატული, ჩეხური, დანიური, ესტონური, ფინური, გალიციური, ბერძნული, ისლანდიური, ირლანდიური, ლატვიური, ლიტვური, მალტური, ნორვეგიული, პორტუგალიური, სერბიული, სლოვაკური, სლოვენური, შვედური, უელსური
მეტყველების დამუშავება ანუ სამეტყველო ურთიერთობა		ინგლისური	ჩეხური, ჰოლანდიური, ფინური, ფრანგული, გერმანული, იტალიური, პორტუგალიური, ესპანური	ბასკური, ბულგარული, კატალონიური, დანიური, ესტონური, გალიციური, ბერძნული, უნგრული, ირლანდიური, ნორვეგიული, პოლონური, სერბიული, სლოვაკური, სლოვენური, შვედური	ხორვატული, ისლანდიური, ლატვიური, ლიტვური, მალტური, რუმინული, უელსური
ტექსტის ანალიზი		ინგლისური	ფრანგული, ჰოლანდიური, გერმანული, იტალიური, ესპანური	ბასკური, ბულგარული, კატალონიური, ჩეხური, დანიური, ფინური, გალიციური, ბერძნული, უნგრული, ნორვეგიული, პოლონური, პორტუგალიური, რუმინული, სლოვაკური, სლოვენური, შვედური	ხორვატული, ესტონური, ისლანდიური, ირლანდიური, ლატვიური, ლიტვური, მალტური, სერბიული, უელსური
ენობრივი რესურსები		ინგლისური	ჩეხური, ჰოლანდიური, ფრანგული, გერმანული, უნგრული, შვედური, იტალიური, პოლონური, ესპანური,	ბასკური, ბულგარული, კატალონიური, ხორვატული, დანიური, ესტონური, ფინური, გალიციური, ბერძნული, ნორვეგიული, პორტუგალიური, რუმინული, სერბიული, სლოვაკური, სლოვენური	ისლანდიური, ირლანდიური, ლატვიური, ლიტვური, მალტური, უელსური

ცხრილი 1 - ევროპის ენების ტექნოლოგიური მხარდაჭერების მახასიათებელი ცხრილი

ქვემოთმოყვანილი ცხრილი_2 ადვილად იგება ცხრილ_1-ში ენების შეფასების სხვადასხვა ჯგუფების ქულებთან გათანაბრებით. კერძოდ, ჩვენ ამ მიზნით: 1. „სრულყოფილი“ გავათანაბრეთ 5 ქულასთან; 2. „კარგი“ – 4-თან; 3. „საშუალო“ – 3-თან; 4. „ფრაგმენტული“ – 2-თან; 5. „სუსტი/არანაირი“ - 1-

თან. ცხრილი_1-დან ასე წარმოქმნილ ცხრილ_2-ში ქართული და ევროპული ენების ტექნოლოგიური მხარდაჭერების შედარებისა და გაანალიზების მიზნით ქართული ენაც ჩავრთეთ. - მცირე განმარტებები იმ მიზნით, რომ ქვემოთმოყვანილი ცხრილი 2 მკითხველისთვის ადვილად გასაგები გახდეს:

1. სვეტში „ენის ჯგუფი და ადგილი“ განთავსებული ორი რიცხობრივი მონაცემიდან პირველი გვიჩვენებს იმ ჯგუფის ნომერს, რომელსაც ენა მიეკუთვნება, მეორე - ამ ენის ადგილს ენათა საერთო რიგში. ამასთან, რაც უფრო დაბალია ენის ჯგუფის ნომერი მით უფრო მაღალია ენის ტექნოლოგიური მხარდაჭერის მაჩვენებელი. ასე მაგალითად, ნორვეგიულ ენას ამ სვეტში უზის მონაცემები - 6 (22). ეს იმას ნიშნავს, რომ ნორვეგიული ენა განეკუთვნება 9 ენისაგან შემდგარ ენათა 6-ე ჯგუფს. ეს კი, თავის მხრივ, იმას ნიშნავს, რომ ნორვეგიული ენის ტექნოლოგიური მხარდაჭერის მაჩვენებელი ნაკლებია უფრო დაბალი ნომრის ჯგუფში განთავსებული ენების ტექნოლოგიური მხარდაჭერის მაჩვენებელზე და მეტია უფრო მაღალი ნომრის ჯგუფში განთავსებული ენების იმავე მაჩვენებელზე. ამასთან, ის, რომ ამ ჯგუფში პირველ ადგილზე რუმინულია, ხოლო ბოლოზე ნორვეგიული, არაარსებითია ანუ ნიშნავს მხოლოდ იმას, რომ ეს ენები სხვა ამ ჯგუფში განთავსებულ ენებთან ერთად ტექნოლოგიური მხარდაჭერის თვალსაზრისით მე-14 - 22-ე ადგილებს იყოფენ 31 ევროპულ ენას შორის, რაც, აქვე ვიტყვით, ამ ენებს ციფრული კვდომის საკმაოდ მაღალი საფეხურის საფრთხის ქვეშ მყოფ ენათა ჯგუფში უჩენს ადგილს. ასე მაგალითად, ნორვეგიულ ენას ამ სვეტში უზის მონაცემები - 6 (22). ეს იმას ნიშნავს, რომ ნორვეგიული ენა განეკუთვნება 9 ენისაგან შემდგარ ენათა 6-ე ჯგუფს. ეს კი, თავის მხრივ, იმას ნიშნავს, რომ ნორვეგიული ენის ტექნოლოგიური მხარდაჭერის მაჩვენებელი ნაკლებია უფრო დაბალი ნომრის ჯგუფში განთავსებული ენების ტექნოლოგიური მხარდაჭერის მაჩვენებელზე და მეტია უფრო მაღალი ნომრის ჯგუფში განთავსებული ენების იმავე მაჩვენებელზე. ამასთან, ის, რომ ამ ჯგუფში პირველ ადგილზე რუმინულია, ხოლო ბოლოზე ნორვეგიული, არაარსებითია ანუ ნიშნავს მხოლოდ იმას,

რომ ეს ენები სხვა ამ ჯგუფში განთავსებულ ენებთან ერთად ტექნოლოგიური მხარდაჭერის თვალსაზრისით მე-14 - 22-ე ადგილებს იყოფენ 31 ევროპულ ენას შორის, რაც, აქვე ვიტყვით, ამ ენებს ციფრული კვდომის საკმაოდ მაღალი საფეხურის საფრთხის ქვეშ მყოფ ენათა ჯგუფში უჩენს ადგილს [13];

ენის ჯგუფი და ადგილი	შესაფასებელი ენა	მეტყველების დამუშავება	მანქანური თარგმანი	ტექსტის ანალიზი	ენობრივი რესურსები	მხარდაჭერის ჯამური ქულა	მხარდაჭერის წილადური და პროცენტული მაჩვენებლები
1 (1)	ინგლისური	4	4	4	4	16 (4)	16/20 = 80% (20%)
2 (2)	ფრანგული	3	3	3	3	12 (8)	12/20 = 60% (40%)
2 (3)	ესპანური	3	3	3	3	12 (8)	12/20 = 60% (40%)
3 (4)	გერმანული	3	2	3	3	11 (9)	11/20 = 55% (45%)
3 (5)	ჰოლანდიური	3	2	3	3	11 (9)	11/20 = 55% (45%)
3 (6)	იტალიური	3	2	3	3	11 (9)	11/20 = 55% (45%)
4 (7)	უნგრული	2	2	2	3	9 (11)	9/20 = 45% (55%)
4 (8)	პოლონური	2	2	2	3	9 (11)	9/20 = 45% (55%)
4 (9)	ჩეხური	3	1	2	3	9 (11)	9/20 = 45% (55%)
4 (10)	გალიციური	2	1	3	3	9 (11)	9/20 = 45% (55%)
5 (11)	კატალონიური	2	2	2	2	8 (12)	8/20 = 40% (60%)
5 (12)	პორტუგალიური	3	1	2	2	8 (12)	8/20 = 40% (60%)
5 (13)	შვედური	2	1	2	3	8 (12)	8/20 = 40% (60%)
6 (14)	რუმინული	1	2	2	2	7 (13)	7/20 = 35% (65%)
6 (15)	სლოვენური	2	1	2	2	7 (13)	7/20 = 35% (65%)
6 (16)	სლოვაკური	2	1	2	2	7 (13)	7/20 = 35% (65%)
6 (17)	ბერძნული	2	1	2	2	7 (13)	7/20 = 35% (65%)
6 (18)	ფინური	2	1	2	2	7 (13)	7/20 = 35% (65%)
6 (19)	დანიური	2	1	2	2	7 (13)	7/20 = 35% (65%)
6 (20)	ბულგარული	2	1	2	2	7 (13)	7/20 = 35% (65%)
6 (21)	ბასკური	2	1	2	2	7 (13)	7/20 = 35% (65%)
6 (22)	ნორვეგიული	2	1	2	2	7 (13)	7/20 = 35% (65%)
7 (23)	ესტონური	2	1	1	2	6 (14)	6/20 = 30% (70%)
7 (24)	სერბიული	2	1	1	2	6 (14)	6/20 = 30% (70%)
8 (25)	ირლანდიური	2	1	1	1	5 (15)	5/20 = 25% (75%)
8 (26)	ხორვატული	1	1	1	2	5 (15)	5/20 = 25% (75%)
9 (27)	ისლანდიური	1	1	1	1	4 (16)	4/20 = 20% (80%)
9 (28)	ლატვიური	1	1	1	1	4 (16)	4/20 = 20% (80%)
9 (29)	ლიტვიური	1	1	1	1	4 (16)	4/20 = 20% (80%)
9 (30)	მალტური	1	1	1	1	4 (16)	4/20 = 20% (80%)
9 (31)	უელსური	1	1	1	1	4 (16)	4/20 = 20% (80%)
10(32)	ქართული	0.5	0.5	0.5	1.5	3 (17)	3/20 = 15% (85%)
11(33)	აფხაზური	0.2	0.2	0.2	0.4	1(19)	1/20 = 5% (95%)

ცხრილი 2 - ქართული და ევროპული ენების ტექნოლოგიური მხარდაჭერების შემადარებელი ცხრილი

2. სვეტში „ტექნოლოგიური მხარდაჭერის ჯამური ქულა“ განთავსებული ორი რიცხობრივი მონაცემიდან პირველი გვიჩვენებს მეტყველების დამუშავების, მანქანური თარგმანის, ტექსტის ანალიზისა და ენობრივი რესურსების მიმართულებით ენისთვის მინიჭებული ქულების ჯამს. ანუ, მოკლედ, ეს რიცხვი გვიჩვენებს ენის ტექნოლოგიური მხარდაჭერის ჯამურ ქულას. ამ სვეტში განთავსებული მეორე რიცხვი გვიჩვენებს იმ ქულას, რამდენიც ენას აკლია სრულყოფილ ტექნოლოგიურ მხარდაჭერამდე. ასე მაგალითად, ჰოლანდიურ ენას ამ სვეტში უზის მონაცემები - 11 (9), რაც იმას ნიშნავს, რომ სრულყოფილ ტექნოლოგიურ მხარდაჭერამდე ჰოლანდიურ ენას აკლია 9 ქულა ამ მიმართულებით უკვე გასრულებული 11 ქულის საპირისპიროდ [13].

3. სვეტში „ტექნოლოგიური მხარდაჭერის წილადური და პროცენტული მაჩვენებელი“ განთავსებული ორი მონაცემიდან პირველი - ტოლობის სახით მოცემული მონაცემი გვიჩვენებს ენის ტექნოლოგიური მხარდაჭერის წილადურ და პროცენტულ მახასიათებლებს, ხოლო მეორე - ფრჩხილებში განთავსებული მონაცემი გვიჩვენებს იმ პროცენტების რაოდენობას, რამდენიც ენას დარჩა სრულყოფილი ტექნოლოგიური მხარდაჭერის მიღწევამდე. ასე მაგალითად, ინგლისურ ენას ამ სვეტში უზის მონაცემები - $16/20 = 80\%$ (20%). ეს იმას ნიშნავს, რომ ინგლისური ენის ტექნოლოგიური მხარდაჭერის წილადური მაჩვენებელია 16/20, პროცენტული კი - 80%. ეს კი, თავის მხრივ, იმას ნიშნავს, რომ სრულყოფილი ტექნოლოგიური მხარდაჭერის უზრუნველყოფამდე ინგლისურ ენას შესასრულებელი დარჩა სამუშაოების 20% უკვე შესრულებული სამუშაოების 80%-იანი წილის საპირისპიროდ [13].

ახლა უკვე, იმის გათვალისწინებით, რომ 2 ცხრილის ყველა სხვა სვეტის შინაარსული დატვირთვა მკითხველისთვის გასაგებია, ქვემოთ, რა თქმა უნდა ქართულ და აფხაზურ ენებზე კონცენტრირებით, ძალიან მოკლედ წარმოვადგენთ ამ უკვე კლასიფიცირებული ცხრილით მოცემული ქართული, აფხაზური და ევროპული ენების ტექნოლოგიური

მხარდაჭერების ურთიერთშედარებითი შეფასების დაპირებული ანალიზის შედეგებს [13]:

1. ციფრული კვდომის საფრთხის ქვეშ მყოფ ენებად არ განიხილება 1-ელ, 2-ე და 3-ე ჯგუფში მოქცეული ენები ანუ, სხვა სიტყვებით, პირველ 6 ადგილზე მყოფი ენები. ეს ის ენებია, რომელთა ტექნოლოგიების პროცენტული მაჩვენებელი 50%-ზე მეტია [13];

2. დანამდვილებით შეიძლება ითქვას, რომ ციფრული კვდომის მეტად მაღალი საფეხურის საფრთხის ქვეშ არიან 5-ე, 6-ე, 7-ე, 8-ე და 9-ე ჯგუფის ენები, ანუ, სხვა სიტყვებით, ის 21 ენა, რომლებიც განთავსებულნი არიან მე-11 – 31-ე ადგილებზე. ამ ენების ტექნოლოგიური მხარდაჭერის პროცენტული მაჩვენებელი თამაშობს 20%-დან 40%-მდე [13];

3. საკმაოდ მაღალი ალბათობით ციფრული კვდომის ასევე მაღალი საფეხურის საფრთხის ქვეშ არიან 4-ე ჯგუფის ენები, ანუ, მე-7 – 10-ე ადგილებზე განთავსებული ენები. ამ ენების ტექნოლოგიური მხარდაჭერის პროცენტული მაჩვენებელია 45% [13].

4. დანამდვილებით შეიძლება ითქვას, რომ 9-ე ჯგუფის ენები ანუ ენები, რომელთა ტექნოლოგიური მხარდაჭერის პროცენტული მაჩვენებელია 20%, ციფრული კვდომის განსაკუთრებით მაღალი საფეხურის საფრთხის ქვეშ არიან [13];

5. დანამდვილებით შეიძლება ითქვას, რომ ქართული ენა, რომელიც მისი 15%-იანი ტექნოლოგიური მხარდაჭერის მაჩვენებლით ჩვენ 10-ე ჯგუფს მივაკუთნეთ, მით უფრო კი აფხაზური ენა, რომელიც მისი 5% ტექნოლოგიური მხარდაჭერის მაჩვენებლით 11-ე ჯგუფის კუთვნილებაა, საგრძნობლად ჩამორჩება 20%-იანი ტექნოლოგიური მხარდაჭერის მქონე 9-ე ჯგუფის ენებს, რაც კიდევ უფრო გასაგებს ხდის, რომ ევროპულ ენებს შორის ყველაზე უფრო მაღალი საფეხურის ციფრული კვდომის საფრთხის ქვეშ მყოფ ისლანდიურ, ლატვიურ, ლიტვიურ, მალტურ და უელსურ ენებთან შედარებით ქართული და აფხაზური ენები კიდევ უფრო მაღალი საფეხურის ციფრული კვდომის საფრთხის ქვეშ იმყოფებიან, რაც კიდევ უფრო ცხადსა

ნათელს ხდის როგორც სადისერტაციო ნაშრომში ზემოთ უკვე ხაზგასმული სამეცნიერო კვლევითი ინსტიტუტის ფორმირების გადაუდებელ აუცილებლობას, ასევე იმას, თუ საიდან მოდის სადისერტაციო კვლევით დასმული მეტად მაღალი ეროვნული მნიშვნელობების მქონე პრობლემა [13].

იმ მიზნით, რომ კიდევ უფრო გასაგები გახდეს რას ნიშნავს და რას გულისხმობს ენის სრული ტექნოლოგიური უზრუნველყოფა და რატომ არის ენა განწირული ციფრული კვდომისთვის, თუ ის ამ ყველა საჭირო საბაზისო ენობრივი ტექნოლოგიით ვერ აღიჭურვა, მოკლედ მიმოვიხილავთ მეტა-ქსელის ტექნოლოგიური საბჭოს ავტორობით გამოქვეყნებულ ნაშრომს „მეტა-ქსელის სტრატეგიული კვლევითი გეგმა 2020 წლის მრავალენოვანი ევროპისათვის“.

ამ გამოცემაში კითხვაზე - „რა არის ენობრივი ტექნოლოგიები?“, მეტა-ქსელის ტექნოლოგიური საბჭო შემდეგ პასუხს იძლევა: *„ენობრივი ტექნოლოგიები არის ტექნოლოგიები, რომლებიც ავტომატურად წარმოქმნიან და, ასევე, ანალიზებენ ჩვენს სამყაროში როგორც წერითი, ისე ზეპირი (ასევე ჟესტიკულაციური) ფორმით სხვადასხვა ენებზე არსებულ ნებისმიერი სპეციფიკისა და ნებისმიერი სირთულის მქონე ინფორმაციის მატარებელ შინაარსებს.“* ანუ, თუ ამ მეტა-ქსელისეული განსაზღვრებიდან გამოვალთ, გასაგები გახდება, რომ ერთი რომელიმე ენის სრული და ამომწურავი აღჭურვა ენობრივი ტექნოლოგიებით ნიშნავს ისეთი კომპიუტერის შექმნას, რომელსაც ეს ენა ანუ ამ ენით მოცემული სააზროვნო და საკომუნიკაციო სისტემა იმდენად კარგად „ეცოდინება“, რომ ნებისმიერი სახის საჭიროების შემთხვევაში ამ ენის მცოდნე ნებისმიერი მომხმარებელი თავისუფლად ანუ შეზღუდვების გარეშე შევა მასთან სრულ და ამომწურავ სასაუბრო სამომხმარებლო ურთიერთობაში. ამგვარად, ამა თუ იმ ენაში ენობრივი ტექნოლოგიებით კომპიუტერის სრული და ამომწურავი აღჭურვა ანუ ამა თუ იმ ენის სრული ტექნოლოგიური უზრუნველყოფა, თავად მეტა-ქსელის ტექნოლოგიური საბჭოს ზემომოყვანილი განსაზღვრებიდან გამომდინარე, ნიშნავს ამ ენის სრულყოფილად მცოდნე კომპიუტერის ანუ

ამ ენის ტექნოლოგიური ანბანის აგებას. თუმცა, მეტა-ქსელის „სტრატეგიული კვლევითი გეგმა 2020 წლის მრავალენოვანი ევროპისათვის“ ანუ ამ ნაშრომის ხედვა არ ჩერდება ერთი ენის მცოდნე კომპიუტერზე და ის უფრო ღრმად მიდის ანუ უკვე საუბრობს „მრავალ ენაზე მოაზროვნე, მოსაუბრე და მთარგმნელ კომპიუტერებს“ და ასეთ „კომპიუტერებზე“ დამყარებულ კულტურულ ეპოქაზე [12]. - ანუ, ამ ყველაფრიდან გამომდინარე ქართული და აფხაზური ენების სრული ტექნოლოგიური უზრუნველყოფა ნიშნავს იმას, რომ სამომავლოდ უკვე დაგეგმილ ენობრივი ბარიერებისაგან თავისუფალ კულტურულ სამყაროში ანუ სამყაროში, *„რომელშიც წერითი და ზეპირი კომუნიკაცია აღარასდროს აღარ შეფერხდება ენობრივი ბარიერებით და რომელშიც არა მხოლოდ ზოგადი, არამედ სპეციალიზებული შინაარსების მაღალი ხარისხის თარგმანებაც საყოველთაოდ ხელმისაწვდომი იქნება“*, ნებისმიერ ქართველ და აფხაზ „მოქალაქეს შეეძლება გამოიყენოს ნებისმიერი მომსახურება, ხელი მიუწვდებოდეს ნებისმიერ ცოდნაზე, ისიამოვნოს ნებისმიერი მედია საშუალებით, გააკონტროლოს ნებისმიერი ტექნოლოგია თავისი მშობლიური ენით“ - ქართული და აფხაზური ენით [12].

ამგვარად, აქ, ზემოთ, იმის თაობაზე დასმულ კითხვას, თუ რა იგულისხმება ენების, მათ შორის, ქართული და აფხაზური ენების ყველა საჭირო საბაზისო ტექნოლოგიით აღჭურვაში და რატომ არის ენა განწირული ციფრული კვდომისთვის, თუ ის ამ ყველა საჭირო საბაზისო ტექნოლოგიით ვერ აღიჭურვა, პასუხს მეტა-ქსელის ზემოციტირებული განცხადებაზე დაყრდნობით ვცემთ. - ანუ, აქ იგულისხმება ის, რომ ჩვენ ანუ ქართველები და აფხაზები ან შევძლებთ „ქართულ და აფხაზურ ენებზე მოაზროვნე, მოსაუბრე და მთარგმნელი მანქანების“ აგებას და ქართული და აფხაზური ენებით შევალთ „მრავალ ენაზე მოაზროვნე, მოსაუბრე და მთარგმნელი მანქანების“ კულტურულ ეპოქაში, ანდა, ეს უკვე კარს მომდგარი ახალი კულტურული ეპოქა კი დადგება, მაგრამ ჩვენ, „ქართულ და აფხაზურ ენებზე მოაზროვნე, მოსაუბრე და მთარგმნელი მანქანების“ არარსებობის გამო, ამ

ახალ ეპოქაში ვერ შევძლებთ ვერც ქართული და ვერც აფხაზური ენით ვერც შესვლასა და ვერც მონაწილეობას, რაც, მთლიანობაში, კიდევ უფრო გასაგებს ხდის, რამ განაპირობა ის, რომ ჩემი სადოქტორო კვლევა მიმართული იყო საქართველოს სახელმწიფო ენების - ქართულისა და აფხაზურის კულტურული დაცვისა და ტექნოლოგიური განვითარების ამოცანებზე [12].

1.4. იმ AR/122/4-105/14 პროექტის შესახებ, რომლის გამაგრძელებელი პროექტიცაა სადისერტაციო კვლევა

როგორც უკვე აღინიშნა, სადოქტორო თემა „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ არის შოთა რუსთაველის ეროვნული სამეცნიერო ფონდის AR/122/4-105/14 წარმატებული პროექტის „კიდევ ერთი ნაბიჯი მოსაუბრე ქართული თვითგანვითარებადი ინტელექტუალური კორპუსისაკენ“ გამაგრძელებელი პროექტი [66-76]. ამასთან, სადოქტორო კვლევა მიზნად ისახავდა AR/122/4-105/14 პროექტის ფარგლებში შემუშავებული ქართული ინტელექტუალური ვებ-კორპუსის, სხვა სიტყვებით, ქართული ჭკვიანი კორპუსის გაძლიერების გზით ქართული უნივერსალური ჭკვიანი კორპუსის აგებას, რისი აგებაც, თავის მხრივ, სრულიად აუცილებელი მოცემულობაა გრძელვადიანი პროექტებით „ქართული ენის ტექნოლოგიური ანბანი“ და „აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის გეგმა-პროგრამა“ დასახული მიზნების მისაღწევად, რაც, როგორც უკვე აღინიშნა, გულისხმობს ქართული და აფხაზური ენების ციფრული კვდომის საფრთხისგან დაცვის მიზნების გათვალისწინებით ქართული და აფხაზური ენების ტექნოლოგიური ანბანებისა და ქართული და აფხაზური ენებში მოაზროვნე, მოსაუბრე და მთარგმნელი მანქანების აგებას. ეს კი, მთლიანობაში, როგორც ასევე უკვე აღინიშნა, მიმართულია ერთიანი სახელმწიფოს ფარგლებში ქართველებისა და აფხაზების ევროკავშირში, უფრო ზოგადად კი, მომავლის კულტურულ

სამყაროში ქართული და აფხაზური ენებით ღირსეული ანუ სრულფასოვანი შესვლის მეტად მაღალი მნიშვნელობების მქონე მიზნების მიღწევაზე.

ამგვარად, კიდევ უფრო გასაგები რომ იყოს, დისერტაციაში დასმული პრობლემა და მისი არსი, ქვემოთ, მოკლედ და ზოგადად არის მიმოხილული AR/122/4-105/14 პროექტით მიღწეულ შედეგები და აქ გამოყენებული მეთოდები, რომლებიც, თავის მხრივ, სრული სახით არის წარმოდგენილი და აღწერილი მონოგრაფიულ ნაშრომში „ქართული ინტელექტუალური ვებ-კორპუსი: მიზნები, მეთოდები, რეკომენდაციები“.

AR/122/4-105/14 პროექტით 2017 წელს აგებული მოსაუბრე ქართული თვითგანვითარებადი ინტელექტუალური კორპუსი, მოკლედ, ქართული ინტელექტუალური ვებ-კორპუსი ანუ, სხვა სიტყვებით, ქართული ჭკვიანი კორპუსი (მისი დღევანდელი გაძლიერებული ვერსია იძებნება <https://corpus.ge/> მისამართზე) მიუხედავად იმისა, რომ ის ჩაფიქრებული იყო ქართული უნივერსალური ჭკვიანი კორპუსის აგების მიმართულებით გადადგმულ ჯერ მხოლოდ პირველ ნაბიჯად, საცდელი სახით:

1. უკვე იყო ქართული ენის ერთადერთი თვითგანვითარებადი ინტელექტუალური კორპუსი. - ის უკვე მაშინ თავადვე აწვითარებდა თავის მოცულობით და ინტელექტუალურ მაჩვენებლებს.

2. უკვე იყო ქართული ენის ერთადერთი ინტერაქტიური კორპუსი. - მასში ჩადგმული სახით უკვე მაშინ ფუნქციონირებდა ხმოვანი მართვის მოდული.

3. უკვე იყო ქართული ენის ერთადერთი მრავალენოვანი კორპუსი. - მასში ჩადგმული სახით უკვე მაშინ ფუნქციონირებდა ქართულ-ინგლისური, ქართულ-რუსული და ქართულ-აფხაზური ვებ-კორპუსები.

4. უკვე იყო ქართული სამეტყველო ენისა და ტიტრირებული სამეტყველო მონაცემების ერთადერთი მრავალმოდულური კორპუსი. - მასში უკვე მაშინ ჩადგმული იყო ქართული სამეტყველო მონაცემებისა და ტიტრირებული სამეტყველო მონაცემების კორპუსის შესაქმნელი ინსტრუმენტები.

5. უკვე იყო ქართული სამწერლობო ენის ყველაზე უფრო მოცულობითი კორპუსი. – აღნიშნულში გვარწმუნებს ქართული ეროვნული კორპუსის (<http://gnc.gov.ge/gnc/static/portal/gnc.html>) მაშინდელი ანუ 2017 წლის მონაცემები, რომელიც უკვე მაშინ თხუთმეტ წელზე მეტია შენების პროცესში იყო და რომელშიც თავმოყრილი იყო მაშინ და დღეს არსებული ქართული კორპუსების მეტი წილი (ესენია: 1. GNC ანუ ქართული ენის ეროვნული კორპუსი (მგელი ქართული), რომლის მოცულობა 4349721 ტოკენია(2016.08.29); 2. GNC ანუ ქართული ენის ეროვნული კორპუსი (საშუალო ქართული), რომლის მოცულობა 1242634 ტოკენია(2016.08.29); 3. GNC ანუ ქართული ენის ეროვნული კორპუსი (ახალი ქართული), რომლის მოცულობა 451584 ტოკენია(2016.08.29); 4. GRC ანუ ქართული ენის რეფერენციალური კორპუსი, რომლის მოცულობა 202728329 ტოკენია(2016.12.05); 5. GDC ანუ ქართული დიალექტური კორპუსი, რომლის მოცულობა 1694362 ტოკენია(2016.09.14); 6. SSGG ანუ სოციოლინგვისტური სიტუაცია თანამედროვე საქართველოში, რომლის მოცულობა 152708 ტოკენია(2016.09.14); 7. GNC ანუ ქართული ენის ეროვნული კორპუსი (მეგრული), რომლის მოცულობა 89 404 ტოკენია(2016.09.14); 8. GNC ანუ ქართული ენის ეროვნული კორპუსი (სვანური), რომლის მოცულობა 473180 ტოკენია(2016.09.14) [13]. - ანუ, ამ მონაცემებით ჩვენი კორპუსი, რომელიც მაშინ ანუ 2017 წელს ჯერ მხოლოდ ორი წლის იყო 255 მილიონ სიტყვაზე მეტს შეიცავდა და ქართული ენის ეროვნულ კორპუსში შემავალი კორპუსების ჯამურ მოცულობას 44 მილიონზე მეტი ტოკენით აღემატებოდა. უფრო სრული სურათის მოცემის მიზნით მოგვყავს აგრეთვე ილიას სახელმწიფო უნივერსიტეტის ქართული ენის კორპუსი, რომელიც 2017 წლის მონაცემებით <http://corpora.iliauni.edu.ge/?q=ka/search-words> მისამართზე იყო განთავსებული და რომელიც, იქმნებოდა რა 2009-2015 წლებში, ჯამურად შეიცავდა 100 მილიონ ტოკენს. - ანუ, ამ შემთხვევაში, ჩვენი იმდროისათვის ორი წლის კორპუსი ამ იმავე დროისათვის შვიდი წლის კორპუსს მოცულობით 156 მილიონ ტოკენზე მეტით აღემატებოდა[13].

გარდა ამისა, ქართულ ჭკვიან კორპუსში უკვე მაშინ საცდელი სახით ფუნქციონირებდა ისეთი უნიკალური ქართული ტექნოლოგიური სისტემები, როგორცაა (ყველა ეს ქვემო ჩამოთვლილი სისტემა დეტალურად არის აღწერილი [10, 13]-ში):

1. ქართული სამწერლობო და სამეტყველო ენის ტექსტების დამამუშავებელი სისტემები;
2. ქართული V, N და A ტიპის სიტყვების ავტომატური ტაგერები/დესკრიპტორები/გენერატორები;
3. ქართული თვითგანვითარებადი სინტაქსური/ორთოგრაფიული მართლმწერი სისტემები;
4. ქართულ-მათემატიკური და ჰიბრიდული ქართულ-ინგლისურ-გერმანული მთარგმნელი სისტემები;
5. ქართული ტექსტების მანალიზებელი, კითხვებზე-მოპასუხე და ლოგიკური ამოცანებისა და ანალოგიების ავტომატური გენერირებისა და ამოხსნის უნარების ავტომატური ტესტირების სისტემები;
6. ქართული შინაარსულად მკითხველი/მსმენელი სისტემები;
7. ქართულ მკითხველ სისტემაში სხვადასხვა სახმო მონაცემების ჩამდები სისტემები;
8. ქართული ტიტრირებული მეტყველების ანალიზატორები/სეგმენტატორები;
9. ქართული ხმოვანი მართვის მოდული;
10. კორპუსის განმავითარებელი ინსტრუმენტის სახით აგებული ქართული ჭკვიანი ინტერნეტ ჟურნალი;
11. მეტყველებაზეზღუდულპირთათვის ქართული სასაუბრო დამხმარე სისტემის სამობილურო/საინტერნეტო აპლიკაციები;
12. ქართული მრავალენოვანი სასაუბრო დამხმარე სისტემის სამობილურო/საინტერნეტო აპლიკაციები;
13. ქართული მრავალენოვანი ხმოვანი ლექსიკონის სამობილურო/საინტერნეტო აპლიკაციები;

14. გუგლის მთარგმნელის ქართული ჰიბრიდული გაფართოების სამობილურო/საინტერნეტო აპლიკაციები;
15. მრავალფუნქციური ქართული ჰკვიანი ფურცელი ანუ ქართული_მოსაუბრე&მოაზროვნე&მთარგმნელი_1 სისტემა და ქართულ-ინგლისურ-გერმანულ-რუსული თვითმასწავლებელი;
16. ქართულ ჰკვიან ფურცლებს შორის მრავალენოვანი ტექსტუალური შეტყობინებების საინტერნეტო/სამობილურო აპლიკაციები;
17. ქართულ ჰკვიან ფურცლებს შორის ხმოვანი შეტყობინებების საინტერნეტო / სამობილურო აპლიკაციები;
18. ქართული ადაპტირებული/ხმოვანმართვიანი, „ჯიმეილი“, ინტერნეტი, ვიკიპედია.

ამგვარად, ეს ყველაფერი, უკვე მაშინ, სრულ საფუძველს იძლეოდა იმის ხაზგასმით თქმისა, რომ ქვეყნისათვის საგრანტო პროექტით მიღწეული შედეგების ეფექტი უნდა შეფასდეს, როგორც ქართული ენის ციფრული კვდომის საფრთხისგან დაცვის ანუ ქართული ენისა და ქართული ენით არსებული კულტურის შენარჩუნების მიმართულებით გადადგმულ მნიშვნელოვან ნაბიჯად, რადგან ტექნოლოგიურად დაწინაურებულ ენებთან შედარებით ენობრივი რესურსებითა და ტექნოლოგიებით მაშინაც საგანგაშოდ ჩამორჩენილმა ქართულმა ენამ პროექტით მიღწეული შედეგების წყალობით ეს ჩამორჩენა საგრძნობლივ შეამცირა. ამის გარდა, ასევე მეტად მნიშვნელოვანი იყო აფხაზური ენის ციფრული კვდომის საფრთხისგან დაცვის მიმართულებით AR/122/4-105/14 პროექტით გადადგმული ის პირველი ნაბიჯი, რომლებიც ცხადია ითხოვდა ამავე მიმართულებით შემდგომი ნაბიჯების გადადგმას.⁵ - აქვე ვიტყვი, რომ ეს შემდეგი ნაბიჯი ჩემი სადოქტორო თემით დაგეგმილი კვლევის ფარგლებში გადაიდგა.

⁵ ეს ყველაფერი, მთლიანობაში, იქცა იმ საფუძვლად, რაზე დაყრდნობითაც შოთა რუსთაველის საქართველოს ეროვნულმა სამეცნიერო ფონდმა ეს AR/122/4-105/14 პროექტი 2017 წლის წარმატებულ პროექტად გამოაცხადა.

ამგვარად, კიდევ ერთხელ ხაზს ვუსვამ, რომ სადოქტორო თემა „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“, რომელიც, როგორც ზემოთაც აღინიშნა, არის ამ ზემოთ მოკლედ მიმოხილული წარმატებული AR/122/4-105/14 პროექტის გამაგრებელი პროექტი, ცხადია სასაფუძვლოდ ეყრდნობოდა ამ AR/122/4-105/14 პროექტის ფარგლებში უკვე მიღწეულ და ზემოთ მოკლედ მიმოხილულ შედეგებსა და, აგრეთვე, ფხაკაძის ქართული ენის ლოგიკურ გრამატიკის მიხედვით შემუშავებულ იმ ფხაკაძე-ჩიქვინიძისა და ფხაკაძე-ჩიჩუას მეთოდებს, რომლებზე დაყრდნობითაც AR/122/4-105/14 პროექტის ეს ზემოაღწერილი შედეგებია მიღწეული.

1.5. აღნიშნათა თეორიისა და ქართული ენის ლოგიკური გრამატიკის ჩვენთვის სასაფუძვლო მნიშვნელობის საკითხები

ამგვარად, ახლა უკვე, იმ მიზნით, რომ კიდევ უფრო გასაგები გახდეს დისერტაციაში გადასაჭრელად დასმული პრობლემისა და ამ პრობლემის გადამწყვეტი მეთოდების არსი, ქვემოთ, ძალიან მოკლედ მიმოვიხილავთ ფხაკაძის ქართული ენის ლოგიკური გრამატიკის იმ ჩვენთვის სასაფუძვლო მნიშვნელობის საკითხებს, რომლებიც სასაფუძვლო იყო აგრეთვე წარმატებული AR/122/4-105/14 პროექტისათვის „კიდევ ერთი ნაბიჯი მოსაუბრე ქართული თვითგანვითარებადი ინტელექტუალური კორპუსისაკენ“, რომელიც ზემოთ მოკლედ უკვე მიმოვიხილეთ.

ქართული ენის ლოგიკური გრამატიკის მიზანია ბუნებრივი ქართული ენით ბუნებრივად მოცემული შინაარსული წესრიგის სრული ახსნა ანუ ამ წესრიგის ბუნებრივად განმსაზღვრელი კანონების ამომწურავი აღწერა და ამ გზით ქართული ენის იზომორფული მათემატიკური თეორიის აგება, რომლის შემდგომი სრული პროგრამული რეალიზაცია ქართული ენის სრულყოფილად მცოდნე ინტელექტუალური კომპიუტერული სისტემის ანუ ქართული ენის ტექნოლოგიური ანბანის აგების საშუალებას მოგვცემს.

ამგვარად, ის, რომ ქართული ენის ლოგიკური გრამატიკა პირდაპირ კავშირშია ხელოვნური ინტელექტის ამოცანით ქართული ენის სრულყოფილად მცოდნე კომპიუტერული სისტემის აგების ამოცანის სრულ და ამომწურავ გადაწყვეტასთან, აქ აღნიშნულით არის განპირობებული.

გარდა ამისა, ძირითადი მიზეზი იმისა, რომ ქართული ენის ლოგიკური გრამატიკით შემოთავაზებული ხედვები არსებითად განსხვავდება იმავე საკითხებზე დღეს არსებული ქართული კლასიკური ხედვებისაგან, გახლავთ ის, რომ ლოგიკური გრამატიკა ქართულს უყურებს მის სიღრმისეულ საფეხურზე მდგარ მათემატიკურ ენასთან ერთად მაშინ როდესაც ქართულ კლასიკურ ენათმეცნიერულ კვლევებში ქართული ენის სიღრმისეულ საფეხურზე მდგარი მათემატიკური ენა მხედველობის მიღმა დარჩენილი და ისინი მხოლოდ ე.წ. „საკუთრივ ლინგვისტური“ პრობლემების კვლევით იყვნენ დაკავებულნი.

ამდენად, აღნიშნულის გათვალისწინებით, გასაგები ხდება, რომ თანამედროვე მათემატიკური - ლოგიკურ-ლინგვისტური თვალსაზრისებით ქართული ენა თითქმის შეუსწავლელია. უფრო მეტიც, თუ არ ჩავთვლით იმ მცირე ძალებით წარმოებულ კვლევით პროცესებს, რომელიც 2012 წლიდან საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის გრძელვადიანი პროექტის „ქართული ენის ტექნოლოგიური ანბანი“ ფარგლებში მიმდინარეობს, ჩვენი მონაცემებით, ქართული ენის თანამედროვე მათემატიკური - ლოგიკურ-ლინგვისტური მეთოდებით შესწავლის სხვა ალტერნატიული კვლევითი პროცესი არც არსად არ წარმოებს.

ასევე, ხაზგასმელია ისიც, რომ ქართული ენის ლოგიკური გრამატიკის საფუძველმდებრთი ცნებები და ამგები მეთოდები, ერთი მხრივ, ითვალისწინებს ჩომსკისა და მონტეგიუს ლინგვისტური ხედვების შემდგომი განვითარების საფუძველზე გამოთვლით ლინგვისტიკაში დღეს უკვე ფორმირებულ კლასიკურ მიდგომებს, მეორე მხრივ კი, არსებითად განსხვავდება მათგან. საქმე ისაა, რომ ქართული ენის ინგლისური ენისაგან

ტიპობრივად განსხვავებულმა სტრუქტურამ, თავიდანვე გასაგები გახადა, რომ ინგლისური სისტემის თავისებურებებზე დაყრდნობით აგებული მეთოდებისა და თეორიების კალკირება ქართულისათვის არ იყო გამართლებული. - აღნიშნულში მკითხველის უფრო ღრმად გარკვევის მიზნით ქვემოთ მოკლედ მიმოვიხილავთ ქართული ენის ლოგიკური გრამატიკით ქართულისთვის შემუშავებულ ახალ ხედვებს.

საზოგადოდ, ბუნებრივი ენის ლოგიკურ გრამატიკა არის ამ ენობრივი სისტემის შემადგენლების მათემატიკურ სიმბოლოებად გადააზრების გზით მიღებული მათემატიკური თეორია. აქედან გამომდინარე, ქართული ენის ლოგიკური გრამატიკა ქართული ენის შემადგენლების მათემატიკურ სიმბოლოებად გადააზრების გზით მიღებული მათემატიკური თეორია. ამასთან, გრძელვადიანი პროექტის „ქართული ენის ტექნოლოგიური ანბანი“ დასკვნითი მიზანი ანუ ქართული ინტელექტუალური და მთარგმნელი სისტემის აგება, ცხადია, უწინარესად აუცილებლობით ითხოვს ქართული ინტელექტუალური სისტემის აგებას, რაც, თავის მხრივ, ასევე უწინარესად, აუცილებლობით ითხოვს ქართული ენით მოცემული სააზროვნო სისტემის ბუნებრივი კანონების სრულად და ზუსტად ამსახველი მათემატიკური თეორიის - ქართული ენის ლოგიკური გრამატიკის შემუშავებას. - ეს ასეა, რადგან ცხადია, რომ ზემოდეფინირებული სახის ქართული ენის ლოგიკური გრამატიკით მოცემული ცოდნის გარეშე არათუ ქართული ინტელექტუალურ სისტემის, არამედ ამ სისტემის ისეთი ქვესისტემის აგებაც კი ვერ მოხერხდება, როგორც არის ვთქვათ ქართული ტექსტების ავტომატური გრამატიკული სინტაქსური მართლმწერი და ანალიზატორი სისტემები.

ამგვარად, ყოველივე ზემოაღნიშნულნი გასაგებს ხდის როგორც იმას, თუ რატომ განიხილება სადისერტაციო პრობლემატიკა და, ასევე, PHDF-18-1228 სადოქტორო პროექტი „ქართული აფხაზური ენებით ევროკავშირში ანუ სადოქტორო თემა - ქართული უნივერსალური ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და

არსებულთა გაუმჯობესება“ საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიზების სასწავლო-სამეცნიერო ცენტრის გრძელვადიანი პროექტის „ქართული ენის ტექნოლოგიური ანბანი“ ერთ-ერთ სასაფუძვლოდ მნიშვნელოვან ქვეპროექტად, ასევე იმასაც, რომ ქართული ენის ლოგიკური გრამატიკა არის ამ ზემოაღნიშნული გრძელვადიანი პროექტისა და სადისერტაციო პრობლემატიკის საერთო მეთოდური - თეორიული საფუძველი [76-86].

აქ გასათვალისწინებელია ისიც, რომ მათემატიკის ფარგლებში, სხვადასხვა ბუნებრივი ენების ტექნოლოგიური დამუშავების მიზნით, ამ ენების ანუ ამ ენებით ბუნებრივად მოცემული სააზროვნო სისტემების მათემატიკური გრამატიკების შემუშავების პროცესები საუკუნეებია თითქმის უწყვეტივ მიმდინარეობს. მართლაც, მიუხედავად იმისა, რომ ენების ტექნოლოგიური დამუშავების მიზნით დღეს დასავლეთში დასკვნით ფაზაში შესული კვლევითი პროცესების თეორიული საფუძვლების შემუშავებაში ჩომსკი და მონტეგიუ პიონერებად, უფრო მეტიც, რევოლუციური მნიშვნელობების მკვლევრებად მიიჩნვიან, მათ, კიდევაც თუ ყურადღებას არ მივაქცევთ ძველი ინდოელებისა და ბერძნების (პანინი, არისტოტელე, და სხვა) გრამატიკულ კვლევებს, მრავალი ევროპული წინამორბედი ჰყავდათ. მათ შორის ისეთი ცნობილი მათემატიკოსები და ფილოსოფოსები როგორც იყვნენ რ.დესკარტი, ბ.პასკალი, გ.ვ.ლაიბნიცი, ჯ.ბული, გ.ფრეგე, ბ.რასელი, ლ.ვიტგენშტეინი, კ.ეჟდუკევიჩი, ე.ლ.პოსტი, ა.ტარსკი, ა.ჩორჩი. ამავე რიგშია ლ.ბლუმფილდი, ოღონდ იმ განსხვავებით, რომ ის იყო ლინგვისტი. თუმცა, სწორედ ბლუმფილდის სტრუქტურალიზმი და პოსტის მაწარმოებელ სისტემები დაედო საფუძვლად ჩომსკის იმ გენერაციული გრამატიკის თეორიას, რომელი თეორიის გამოც ჩომსკის კვლევები შეფასდა რევოლუციურად. რაც შეეხება მონტეგიუსეულ ინგლისური ენის ლოგიკურ გრამატიკას, აქ მისი ძირითადი საფუძვლები იყო სემანტიკის კომპოზიციურობის ფრეგესეული პრინციპი. ამასთან, ის, რომ მისი კვლევები რევოლუციურად ითვლება, განპირობდა იმით, რომ

ჩომსკისგან განსხვავებით, მონტეგიუ ლინგვისტური კვლევებისას სიმრავლური სემანტიკების ფარგლებში სემანტიკურ ასპექტებსაც განიხილავს.

ზემოთ, ძალიან მოკლედ წარმოვადგინეთ 1950-1970 წლებში მიმდინარე ჩომსკისა და მონტეგიუს რევოლუციური ხასიათის ლინგვისტური კვლევების ისტორია და საწყისები. მას შემდეგ ეს კვლევები მეტად ინტენსიურად განვითარდა. ამ პროცესებმა ბუნებრივად განაპირობა ახალი ბიძგები და დაიწყო აგრეთვე ახალი, რიგ შემთხვევებში გარკვეული აზრით ალტერნატიული სახის მათემატიკური ანუ ფორმალური გრამატიკების შემუშავების საკმაოდ ინტენსიური პროცესები, რისი მაგალითებიცაა: რჰაუსერის გაგრძელებადი გრამატიკა; კ.პოლარდისა და ი.ზაგის თავით მართული სტრუქტურათა გრამატიკა, რომლის უშუალო წინაპარია ჯ.გაზდარის განზოგადებულ ფრაზათა სტრუქტურული გრამატიკა; დ.პარლმუტერისა და პ.პოსტალის რელაციური გრამატიკა; ს.კორნელის ფუნქციონალური გრამატიკა და მისი მემკვიდრე ფუნქციონალურ დისკურსული გრამატიკა; მ.ჰოლიდის სისტემატური ფუნქციონალური გრამატიკა და სხვა მრავალი. - ამით ჩვენ იმის თქმა გვინდა, რომ, დღეს, ერთი მხრივ, ტექნოლოგიურად დაწინაურებული ენები, მიუხედავად იმისა, რომ ისინი უკვე 50 წელზე მეტია ინტენსიურად შეისწავლება თანამედროვე მათემატიკური მიდგომებით, კვლავაც მზარდი ინტენსივობით მიმდინარე მათემატიკური კვლევების ქვეშ არიან, მეორე მხრივ კი, თუ არ ჩავთვლით FR/362/4-105/12 პროექტით „ქართული ენის ლოგიკური გრამატიკის საფუძვლები და მისი გამოყენება საინფორმაციო ტექნოლოგიებში“ შემუშავებულ ქართული ენის ლოგიკური გრამატიკის საფუძვლებს, ჯერ კიდევ შეთანხმებული სამეცნიერო ცოდნის სახით არ არსებობს ქართული ენის რაიმე სხვა ალტერნატიული მათემატიკური თეორია [24]. - ეს და ამ ფონზე ქართული ენის ტექნოლოგიური უზრუნველყოფის საგანგაშოდ დაბალი მოცულობა და ხარისხი, კიდევ ერთხელ გასაგებს ხდის ქართული ენის მათემატიკური შესწავლის შემდგომი

გაძლიერების გადაუდებელ აუცილებლობას, რადგან სრულიად, რომ ამის გარეშე ქართული ენის ციფრული კვდომის საფრთხისგან დაცვის მიზნებში სწრაფი წინსვლები არ გვექნება.⁶

როგორც უკვე ვნახეთ ბუნებრივი ენების სრული ტექნოლოგიური დამუშავების მიზნით წარმოებულ კვლევებში პიონერები იყვნენ ნ.ჩომსკი და რ.მონტეგიუ, რომლებმაც 1950-1970 წლებში საფუძვლები ჩაუყარეს დღეს ჩომსკი-მონტეგიუს სახელებით ცნობილ მათემატიკურ ლინგვისტურ მიმართულებებს, რითაც საფუძვლები ჩაუყარეს ბუნებრივი ენების სრული ტექნოლოგიური დამუშავების დღესაც ინტესიურად მიმდინარე კვლევებს. ამასთან, ვნახეთ ისიც, რომ მიუხედავად მათი ამ პიონერული როლისა, ისინი ეყრდნობოდნენ მაშინ მყარად ფორმირებული დასავლური მათემატიკური ლოგიკის ფარგლებში უკვე შემუშავებულ თეორიებს. კერძოდ, ჩომსკისთვის ასეთი სასაფუძვლო თეორია იყო ე.ლ.პოსტის მაწარმოებელი სისტემები, მონტეგიუ კი გარდა ფრეგეს კომპოზიციურობის პრინციპის გარდა დაეყრდნო ასევე ჩორჩის λ ოპერატორს, თერმთა გადამწერ სისტემებსა და მაშინ უკვე შემუშავებულ სიმრავლურ ინტერპრეტაციებსა და სემანტიკებს. ანალოგიური ვითარებაა კ.ფხაკაძის ქართული ენის ლოგიკურ გრამატიკის შემთხვევაში, რომელიც მათემატიკის მხრიდან ეყრდნობა შ.ფხაკაძის აღნიშვნათა თეორიასა და ამ თეორიის ფარგლებში მის მიერვე შემუშავებულ მათემატიკის დაფუძნების სემანტიკურ პროგრამას.⁷

⁶ იმის გათვალისწინებით, რომ აფხაზური ენის ბუნებრივად მცოდნე აფხაზი მკვლევარების ჩაურთველად ჩვენ ვერ შევძლებთ აფხაზური ენის მათემატიკური თეორიის შემუშავებას, ხაზს ვუსვამთ იმას, რომ „აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის გეგმა-პროგრამის“ მთავარი სამოქმედო ხაზია პროექტით მიმდინარე კვლევებში აფხაზი მკვლევარების საკმარისი ჩართულობის დროული უზრუნველყოფა.

⁷ იგულისხმება საქართველოში კლასიკური მათემატიკური ლოგიკის სკოლის ფუძემდებელი, მეცნიერებათა დამსახურებული მოღვაწე, შალვა ფხაკაძე - იხილეთ <http://www.nplg.gov.ge/bios/ka/00009469/> ; <http://gmu.ge/index.php?page=pkhakadze&lang=ge> და <https://geoanbani.com/TbiLLAI> მისმართებზე განთავსებული მასალები.

ამგვარად, აქ უკვე აღნიშნულის გათვალისწინებით, ძალიან მოკლედ თავად ამ თეორიაზე: შ.ფხაკაძის აღნიშვნათა თეორიაში დაძებნილია შემამოკლებელ ფორმებზე ოპერირების ისეთი რაციონალური წესთა სისტემა, რომ მათი დახმარებით, ერთი მხრივ, შეიძლება შემოვიტანოთ კლასიკურ მათემატიკურ თეორიებში გამოყენებადი ყველა შესაძლო სახის შემამოკლებელი სიმბოლო და, მეორე მხრივ, ამ წესებით შემოტანილი შემამოკლებელი ფორმები ფლობენ ისეთ თვისებებს, რომლებიც უზრუნველყოფენ ამ ფორმებზე ოპერირების ისეთ თავისუფლებას, რომ შესაძლებელი ხდება ნებისმიერი მათემატიკური მსჯელობის ძირითადი თეორიიდან ახალი სიმბოლოებით გაფართოებულ თეორიებში გადატანა.

გარდა ამისა, შ.ფხაკაძის ახალი შემამოკლებელი სიმბოლოს განსასაზღვრავად საკმარისია საწყისი თეორიის სიმბოლოების შინაარსების განსაზღვრა. ასევე, შ.ფხაკაძის შემამოკლებელი სიმბოლოები თეორიაში არ ცხადდება აქსიომებად და, ამასთან, მათი განმსაზღვრელი წესებით რეალიზდება გაფართოებული თეორიის ფორმების ძირითად თეორიის ფორმებზე დაყვანის დასრულებადობა, რაც ავტომატურად უზრუნველყოფს შ.ფხაკაძისეული გაფართოების კონსერვატიულობას [83-85].

ეს ყველაფერი გასაგებს ხდის როგორც საკითხის შ.ფხაკაძისეული მიდგომის მეთოდურ მნიშვნელობას, ისე მის პრინციპულ განსხვავებულობას მანამდე არსებული მიდგომებისგან. კერძოდ, შ.ფხაკაძის შემამოკლებელი სიმბოლოებისა და ამ სიმბოლოებით წარმოქმნილი გაფართოებული თეორიების ეს ზემოწარმოდგენილი მახასიათებლები ანუ ის, რომ ამ ყველაფრით გაფართოებული თეორიის ინტერპრეტირებადობის ანუ არაწინააღმდეგობრიობის საკითხი ავტომატურად იხსნება დღის წესრიგიდან და, შესაბამისად, სრულად დაიყვანება საწყისი თეორიის ინტერპრეტირებადობის ანუ არაწინააღმდეგობრიობის საკითხზე, არსებითია იმით, რომ ეს იძლევა მცირე მოცულობის ალფაბეტის მქონე არაწინააღმდეგობრივ თეორიაზე დაყრდნობით დიდი მოცულობის არაწინააღმდეგობრივი თეორიების მარტივად ანუ ავტომატურად აგების

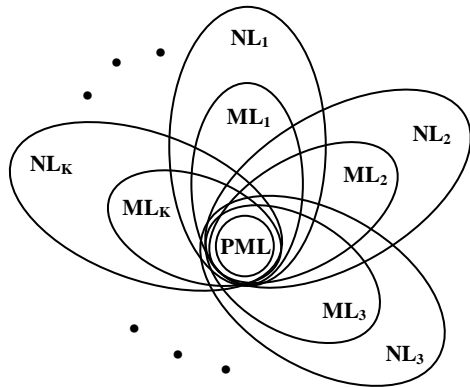
საშუალებას. ეს კი, თავის ხმრივ, მეტად მნიშვნელოვანი შედეგია ბუნებრივი ენების მათემატიკური შესწავლის თვალსაზრისებით.

მართლაც, როგორც ვნახეთ, აღნიშნათა თეორია არის ფორმალური თეორიების განვითარების შ.ფხაკაძისეულ წესთა სისტემა. ამასთან, ფორმალურ თეორიას უწოდებენ განვითარებადს, შესაბამისად არაგანვითარებადს იმის მიხედვით, არის თუ არა მასში რეალიზებული შ.ფხაკაძის განვითარებელ წესთა ტიპის წესები. ანუ, იმის გათვალისწინებით, რომ ფრეგე-ჰილბერტისეული მათემატიკური თეორია ამავდროულად ფხაკაძისეული \mathcal{N} საკმარისად ზოგადი მათემატიკური თეორიაა, შეიძლება ითქვას, რომ აღნიშნათა თეორია მანამდე არაგანვითარებად ფრეგე-ჰილბერტისეულ ფორმალურ თეორიებს განვითარებად თეორიებად აყალიბებს.

ასევე, თუ \mathcal{N}^* არის ნებისმიერი გაფართოება ნებისმიერი არაწინააღმდეგობრივი \mathcal{N} თეორიის და თუ ეს გაფართოება გაკეთებულია შ.ფხაკაძის შემამოკლებელი წესების ტიპის წესებით, მაშინ \mathcal{N}^* აგრეთვე არის არაწინააღმდეგობრივი თეორია.

აქ მოკლედაა გაშუქებული მათემატიკის დაფუძნების შ.ფხაკაძისეული სემანტიკური პროგრამა, რომელსაც მათემატიკის სემანტიკური დაფუძნების პროგრამასაც უწოდებენ ხოლმე.

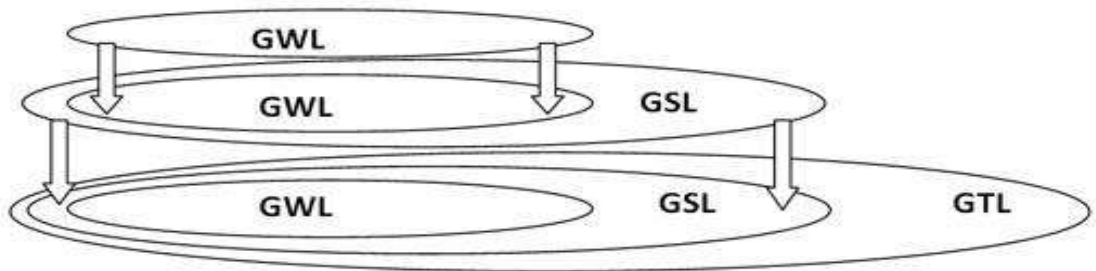
უნდა აღინიშნოს, რომ აღნიშნათა თეორიის ენობრივ სისტემებზე მიყენების პირველი წინჭვრეტა კ.ფხაკაძესთან მათემატიკის დაფუძნების ზემოთ წარმოდგენილმა პროგრამამ განაპირობა. საქმე ისაა, რომ აღნიშნათა თეორია, რომელიც მცირე მოცულობის ალფაბეტით აგებულ არაწინააღმდეგობრივ თეორიაზე დაყრდნობით იძლევა დიდი მოცულობის არაწინააღმდეგობრივი თეორიის აგების მარტივ საშუალებებს, მაღალ სავარაუდოა, რომ იძლეოდეს აგრეთვე ისეთი დიდი მოცულობის მქონე თეორიის ამგები წესების ახსნის გზებს, როგორცაა სხვადასხვა ენებით



მოცემული სხვადასხვა ენობრივი სისტემები. ეს ახალი ხედვა მერე და მერე კიდევ მეტად გამყარდა ქართული და მათემატიკური ენების კ.ფხაკაძის მიერ დასაბუთებული ზოგადი ტიპობრივი ერთგვარობით, რაც, მთლიანობაში, ბუნებრივი

სურათი 1 - ენობრივი ყვავილი მათემატიკური ჰულით

ენების გენეზისის თაობაზე შემდეგი ხედვის სახით ჩამოყალიბდა: ნებისმიერი ბუნებრივი ენა არის პირველადი მათემატიკური ენის გაფართოების შედეგი. - ანუ, ამ მიდგომით პირველადი მათემატიკური ენა ის ბუნებით თანდაყოლილი მოცემულობაა, მხოლოდ რომლის საფუძველზეც ადამიანში აიგება და გაიგება როგორც ქართული, ისე სხვა ნებისმიერი ბუნებრივი ენა. - ამ ჩვენთვის სასაფუძვლო ხედვის მეტი სისრულით წარმოჩენის მიზნით, იხილეთ ქვემოთ განთავსებული სურათი 2.



სურათი 2 - ქართული ენა მის წერით, ზეპირ და სააზროვნო შრეებში

სურათ_2-ზე: 1. GTL-ით აღნიშნულია ქართული სააზროვნო ენა; 2. GSL-ით - ქართული ზეპირი ენა; 3. GWL-ით - ქართული წერითი ენა. ამასთან, როგორც ეს სურათ_2-იდანაც ჩანს, ქართული სამწერლო ენა ჩადგმულია ქართული სამეტყველო ენაში, რომელიც თავის მხრივ, ჩადგმულია ქართულ სააზროვნო ენაში. - უფრო გასაგები რომ იყოს, სხვაგვარად ზუსტად იგივე, რაც ზემოთ ითქვა შემდეგნაირად ყალიბდენა: ყველაფერი რაც ითქმის, არ

იწერება, რაც იფიქრება, არ ითქმება. - მეთოდოლოგიურად ეს იმას ნიშნავს, რომ ჩვენ ვცდილობთ რა ქართული წერითი ენის მართლმწერი და მანანალიზებელი სისტემების აგებას, ქართული წერითი ენის სიტყვებს განვიხილავთ ქართულ აზრით ანუ მათემატიკურ ენაში განსაზღვრულ შ.ფხაკაძისეულ შემამოკლებელ სიმბოლოებად, რითაც, ჩვენი მიდგომები განსხვავდება ლინგვისტურ კვლევებში მანამდე გამოყენებული მიდგომებისგან, რაც, საბოლოო ჯამში, დამატებითი დასტურია ამ ახალ ხედვაზე დამყარებული სადისერტაციო კვლევის მეცნიერული სიახლისა. ანუ, მიუხედავად იმისა, რომ ქართული მრავალი სხვადასხვა მიდგომით უკვე მრავალი წელია შეისწავლება, არცერთი ეს კვლევა არ ითვალისწინებდა ქართული და მათემატიკური ენების ბუნებრივ და თანდაყოლილ ურთიერთკავშირებს.

ამგვარად, ქართული ენის მათემატიკური თეორიის შემუშავების მიზნით პირველი ნაბიჯები დიდი დაგვიანებით - 2000 - 2005 წლებში გადაიდგა და ამ კვლევების შედეგად კ.ფხაკაძემ შეიმუშავა ქართული ენის ლოგიკური გრამატიკის რიგი ფუნდამენტური ცნებები და მეთოდები, რის საფუძველზეც აღწერა ქართული ენის ძირეული ნაწილის მათემატიკური თეორია, რითაც მყარი საფუძველი დაედო ქართული ენის პირველი მათემატიკური თეორიის - ქართული ენის ლოგიკური გრამატიკის შემუშავების დღეს მიმდინარე პროცესებს. ანუ, ზემოაღნიშნულთა გათვალისწინებით, ქვემოთ, მოკლედ გავაშუქებთ კ.ფხაკაძის ქართული ენის ლოგიკური გრამატიკის რიგ ფუნდამენტურ საკითხებს, რომლებიც ასევე ფუნდამენტური მნიშვნელობისაა წინამდებარე სადისერტაციო კვლევისთვის [87].

მათემატიკური და პროგრამული ენების თეორიაში და, ასევე, ბუნებრივ ენებში, სემანტიკური ინტერესების გათვალისწინებით გამოიყოფა დეკლარაციული ანუ თხრობითი, პროცედურული ანუ ბრძანებითი და ინტეროგაციული ანუ კითხვითი წინადადებებათა სამი პრინციპულად განსხვავებული ტიპი. ამასთან, გაგებადობის ზოგადიერარქიული

თვალსაზრისებიდან გამომდინარე, მიიჩნევა, რომ ამ სამი ტიპიდან სასაფუძვლოა პირველი. ამგვარად, როცა განიხილება ბუნებრივი ენობრივი სისტემებისთვის ისეთი ფუნდამენტური დატვირთვის მქონე საკითხი, როგორც არის წინადადების პრობლემა, პირველგადასაწყვეტი ამოცანა ცხადია არის თხრობითი წინადადების პრობლემა. ამასთან, კლასიკური ენათმეცნიერული მიდგომებით *“თხრობითია წინადადება, რომელიც გადმოგვცემს რაიმე ცნობას ამ თუ იმ ფაქტის ან მოვლენის შესახებ, ადასტურებს რასმეს ან უარყოფს”*. – ეს განსაზღვრება არის როგორც მისაღები, ისე გასაგები, მაგრამ კ.ფხაკაძე ქართული ენის ლოგიკურ გრამატიკაში ამაზე არ ჩერდება. ანუ, [52]-ში მან საკითხი შემდეგნაირად დააყენა: როგორ ახერხებს თხრობითი წინადადება რაიმე ცნობის ანუ შეტყობინების გადმოცემას? - ანუ, როგორ ახერხებს იგი რაიმე ფაქტის ან მოვლენის ან დადასტურებას, ან უარყოფას? - კითხვაზე პასუხის გაცემის მიზნით განვიხილოთ წინადადებები ‘წერს.’ ‘და ‘არ წერს.’ და დავსვათ კითხვა: რა შეტყობინებას იძლევა ეს წინადადებები და როგორ? -ამგვარად:

1. თუ ჩვენ ვიცით, რომ ‘წერს’ წინადადება ითქვა როგორც ჭეშმარიტი, მაშინ ამ წინადადებით იგივე შეტყობინება მოგვეწოდება, რაც მოგვეწოდება, როცა ‘არ წერს.’ წინადადება ითქმებოდა როგორც მცდარი;

2. თუ ჩვენ ვიცით, რომ ‘წერს’ წინადადება ითქვა როგორც მცდარი, მაშინ ჩვენ ამ წინადადებით იგივე შეტყობინება მოგვეწოდება, რაც მოგვეწოდება, როცა ‘არ წერს.’ წინადადება ითქმებოდა როგორც ჭეშმარიტი;

3. თუ მეტყველი პირის მიერ ‘წერს’ ან/და ‘არ წერს’ წინადადების თქმისას არ ვიცით, ის ამას ამბობს როგორც ჭეშმარიტი თუ როგორც მცდარი, მაშინ, ცხადია, ამგვარი თქმით ჩვენ არ მოგვეწოდება არანაირი შეტყობინება. - ეს იმას ნიშნავს, რომ ჭეშმარიტულად შეუფასებელი თქმა ვერ იქნება თხრობითი წინადადება.

ანუ, მეტყველი პირის მიერ გამოთქმული რაიმე მოცემულობა მსმენელი პირისთვის მხოლოდ იმ შემთხვევაშია თხრობითი წინადადება, თუ იგი ამ მოცემულობასთან ერთად მეტყველი პირისგან იღებს მონაცემს

იმის შესახებაც, ითქვა ის როგორც ჭეშმარიტი თუ როგორც მცდარი მოცემულობა. - ანუ, მეტყველ პირს არა თუ უფლება, არამედ არანაირი თეორიული საშუალებაც კი არ აქვს იმისა, რომ გამოთქვას თხრობითი წინადადება მასზედ მისგან განუყოფლად მიბმული ლოგიკური დამოკიდებულების ანუ ჭეშმარიტული შეფასების გარეშე. - ამ ზემოწარმოდგენილი მსჯელობით [52]-ში კ.ფხაკაძემ, დაასაბუთა, რომ ნებისმიერი S თხრობითი წინადადება გამოითქმება როგორც ჭეშმარიტული მნიშვნელობით შეფასებული ანუ მარკირებული მოცემულობა და რომ ეს ნებისმიერი S თხრობითი წინადადება არის [- = t] სახის 1-ადგილიანი ოპერატორით აგებული [S = t] სახის აზრობრივი ენობრივ-ლოგიკური მონაცემი.⁸

ამგვარად, ზემოწარმოდგენილი მსჯელობით დასაბუთდა არა მხოლოდ თხრობითი კილოს ოპერატორის არსებობა, არამედ ისიც, რომ ეს ოპერატორი ნებისმიერ ბუნებრივ ენაშია, რადგან დასაბუთდა, რომ მის გარეშე ვერცერთ ენაში ვერცერთ შინაარსის მატარებელ თხრობით მოცემულობას ანუ თხრობით წინადადებას ვერ გამოვთქვამდით. ანუ: თუ ეს თხრობითი კილოს [- = t] ოპერატორი, ერთი მხრივ, ნებისმიერ ენაშია, მეორე მხრივ კი, ცხადი სახით არცერთში, სად არის იგი?⁹ - კითხვაზე პასუხის ძიებამ კ.ფხაკაძე აიძულა ენაში გამოეყო ერთმანეთთან უწყვეტ კავშირში მყოფი ცნობიერი და ქვეცნობიერი შრეები. ანუ, ამ მიდგომით, ნებისმიერი ენა იშლება ცნობიერ და ქვეცნობიერ შრეებად, რომლებსაც ერთმანეთთან შ.ფხაკაძის შემამოკლებელი წესების მაგვარი წესებით ბუნებრივად რეალიზებული კავშირები აქვთ. - ანუ, ზემოთ უკვე დასაბუთდა თხრობითი

⁸ კ.ფხაკაძე დამაჯერებლად არ მიიჩნევს იმას, რომ ენაში თხრობითი წინადადება მცდარი ჭეშმარიტული მნიშვნელობით მოიცემოდეს. თუმცა, ქართულში (და არა მხოლოდ ქართულში), ე.წ. ირონიული თქმისას, ნათქვამი სწორედ [- = f] ოპერატორით მოიცემა, რაც გასაგებს ხდის, რომ [- = f] ოპერატორს ირონიული თქმის ოპერატორი უნდა ეწოდოს.

⁹ ეს კითხვა, ჩვენთვის მნიშვნელოვანი კითხვაა, რადგან ჩვენი პროგრამისტული მიდგომები სწორედ ამ კითხვის პასუხით განსაზღვრული ენობრივ ხედვას ემყარება.

კილოს [- =:t] ოპერატორის არსებობა და ისიც, რომ თხრობითი წინადადების თხრობითობა უზრუნველყოფილია სწორედ ამ ყველა ენაში ქვეცნობიერი სახით აზრით ენაში არსებული თხრობითი კილოს ოპერატორით. ანუ, ჩვენ, სამეტყველო ენის თხრობითი გამოთქმა, რომელიც სამწერლებო ენაში ‘.’-ით აღინიშნება, უკვე გადავიაზრეთ როგორც შემამოკლებელი სიმბოლო ანუ უკვე დავასაბუთეთ, რომ ნებისმიერ ენაში ნებისმიერი თხრობითი წინადადება, როგორც შემამოკლებელი გამოსახულება აღნიშნავს აზრითი ენის [[S] = t] სახის გამოსახულებას, რომელიც აგებულია აზრით ენაში არსებული თხრობითი კილოს მარცხნივ ერთადგილიანი [- =: t] ოპერატორით. - ამდენად, მნიშვნელოვანი დასკვნის სახით შეიძლება ითქვას, რომ ნებისმიერი ბუნებრივი ენის მათემატიკური შესწავლა და, შესაბამისად, ტექნოლოგიური დამუშავება ვერ იქნება სრულყოფილი და წარმატებული თუ ამ ბუნებრივ სამეტყველო ენასთან ერთად არ განვიხილავთ ქვეცნობირ შრეში მასთან უწყვეტ ბუნებრივ აზრობრივ კავშირში მყოფ აზრით ენას ანუ, როგორც ჩვენ მას ზოგჯერ ვუწოდებთ ხოლმე, ბუნებრივ მათემატიკურ ენას.

თხრობით წინადადებაზე ასეთი ხედვა ქართული ენის ლოგიკურ გრამატიკაში [52] იძლევა თხრობითი წინადადების კ.ფხაკამისეულ ბუნებრივ განსაზღვრებას: ამა თუ იმ ენის ესა თუ ის ენობრივი გამოსახულება არის თხრობითი წინადადება, მაშინ და მხოლოდ მაშინ, თუ ამ ენობრივი გამოსახულებით მოწოდებულ შინაარსზე კითხვის დასმა იმის თაობაზე, იგი ჭეშმარიტია თუ მცდარი, აზრს მოკლებული არ არის. - უფრო გასაგები რომ იყოს ეს განსაზღვრება განვიხილოთ შემდეგი ორი გამოსახულება: ‘ეს სკამი’ და ‘ეს სკამია’. თხრობითი წინადადების ზემოშემოთავაზებული კ.ფხაკამისეული განსაზღვრებიდან გამომდინარე: 1. ‘ეს სკამი’ არ არის თხრობითი წინადადება, იმის გამო, რომ მისით მოწოდებულ შინაარსზე კითხვა - ჭეშმარიტია იგი თუ მცდარი? - არ დაისმის; 2. ‘ეს სკამია’ არის თხრობითი წინადადება, იმის გამო, რომ მისით მოწოდებულ შინაარსზე კითხვა - ჭეშმარიტია იგი თუ მცდარი? - დაისმის. - ერთ-ერთი მიზეზი იმისა, რომ ამ განსაზღვრებას ბუნებრივს ვუწოდებთ, ისაა, რომ ქართული ენის

ბუნებრივად მცოდნენი - მათ შორის დაბალი კლასის მოსწავლეებიც - ყოველგვარი დამატებითი გონებრივი ძალისხმევების გარეშე ითვისებენ მას და, შემდგომ უკვე, მასზე დაყრდნობით, უშეცდომოთ ახერხებენ ენაში ამ გაგებით თხრობითი წინადადებების გამოყოფას.

ამასთან, ქართულში, იმით, რომ ნებისმიერი V ტიპის სიტყვით ანუ თხრობითი კილოს ზმნით მოცემულ შინაარსზე კითხვა - არის თუ არა ის ქვეშარიტი? - დაისმის, დასტურდება ის, რომ ქართული თხრობითი კილოს ზმნა ამავედროულად წინადადებაა, რომელიც ქართულ სააზროვნო ანუ მათემატიკურ ენაში $[V = t]$ სახით გადაიწერება. - ის, რომ ქართული თხრობითი კილოს ზმნა ამავედროულად არის წინადადება, საქვაკუთხედო მოცემულობაა როგორც კ.ფხაკაძის ქართული ენის ლოგიკური გრამატიკისთვის, ასევე ჩვენი პროგრამისტული მიდგომებისათვის.

გარდა ამისა, [52]-ში განვითარებული მიდგომებით ქართული თხრობითი კილოს ზმნა იმდენ ადგილიანია, რამდენი განსხვავებული დასრულებული შინაარსის სიტყვით მოცემული გაგებაც არის აუცილებელი და საკმარისი მისი აზრითი შინაარსის გასაგებად. ასევე, ერთი ზმნით აგებული წინადადების იმ წევრს, რომელიც აუცილებელია ამ მარტივი წინადადების ამგები ზმნის შინაარსის გასაგებად, წინადადების აუცილებელ წევრს ვუწოდებთ.

ამ მიდგომებით ანუ ლოგიკური გრამატიკით არსებითი სახელის, ანუ N ტიპის სახელის ესა თუ ის $N\alpha$ მორფოლოგიური ფორმა ბრუნვითი ფორმაა, მაშინ, და მხოლოდ მაშინ, თუ არსებობს ერთი მაინც ისეთი ზმნა და ამ ზმნით აგებადი ერთი მაინც ისეთი მარტივი წინადადება, რომლის ერთი მაინც აუცილებელი წევრი $N\alpha$ მორფოლოგიური ფორმითაა მოცემული. - ამ განსაზღვრებაზე დაყრდნობით კ.ფხაკაძემ დაასაბუთა ქართულში 14 სხვადასხვა ბრუნვითი ფორმის არსებობა [52]. ქვემოთ, ძალიან მოკლედ, დალაგებული და კლასიფიცირებული სახით წარმოდგენილია ქართულში არსებული ეს 14 ბრუნვა და ასევე ის ენობრივი მაგალითები, რომლებითაც მტკიცდება ამა თუ იმ მორფოლოგიური ფორმის კ.ფხაკაძის ლოგიკურ

ბრუნებაში ჩართულობის აუცილებლობა:¹⁰

№ ბრუნვები:

- 1.1 ანუ 1: № ბრუნვა: ვაშლი, გაუცვალა(№ა, №ს, №ი, №ში)
- 1.2 ანუ 2: №ს ბრუნვა: დათვის ეშინია (№ს, №ის)
- 1.3 ანუ 3: №ისთვის ბრუნვა: დისთვის უჩუქებია(№ს, №ისთვის, №ი)
- 1.4 ანუ 4: №ისკენ ბრუნვა: ტყისკენ გადაუხვია(№ა, №იდან, №ისკენ)
- 1.5 ანუ 5: №ისგან ბრუნვა: ქვისგან გამოაქანდაკა (№ა, №ი, №ისგან)
- 1.6 ანუ 6: №იდან ბრუნვა: ხიდან ჩამოვარდა (№ი, №იდან)
- 1.7 ანუ 7; №ით ბრუნვა: ბადით მოიხიბლა (№ი, №ით)

№ს ბრუნვები:

- 2.1 ანუ 8: №ს ბრუნვა: შვილს უყვარს(№ს, №ი)
- 2.2 ანუ 9: №ში ბრუნვა: ვაშლში გაუცვალა(№ა, №ს, №ი, №ში)
- 2.3 ანუ 10: №ზე ბრუნვა: მაგიდაზე დადო(№ა, №ი, №ზე)
- 2.4 ანუ 11: №ად ბრუნვა: ვეფხვად გადააქცია(№ი, №ად)
- 2.5 ანუ 12: №ამდე ბრუნვა: ხემდე გააბა(№ა, №ი, №იდან, №ამდე)
- 2.6 ანუ 13: №თან ბრუნვა: მაგიდასთან მიგორდა(№ი, №თან)

№ა ბრუნვა:

- 3.1 №ა ანუ 14: ბრუნვა: კაცმა მიაგორა (№ა, №ი, №თან)

ცხრილი 3 - ლოგიკური ბრუნების ცხრილი

ბრუნების ამ ზემოწარმოდგენილ სისტემას ლოგიკურს ვუწოდებთ, რადგან ქართული ენაში ნებისმიერი არსებითი სახელის ნებისმიერი №ა ბრუნვითი ფორმა 1-ადგილიანი ოპერატორია, რომლითაც მოკლდება ქართულ მათემატიკურ ენაში მისით მოცემული მარტივი პირდაპირი ჩასმის №ა ოპერატორი. - ამ ჩვენთვის სასაფუძვლო ახალი ხედვის კიდევ უფრო ღრმად გასააზრებლად განვიხილოთ ერთი მაგალითი. ამ მიზნით ჯერ მიმოვიხილოთ შემდეგი სამაგალითოთ საჭირო სიტყვები: 1. გაუცვალა,

¹⁰ ქართულში ბრუნების კ.ფხაკამისეულ და ტ.ფუტკარამისეულ გააზრებებს შორის დიდი ფორმოზრვი მსგავსებაა. თუმცა, ამოსავალ სასაყრდენე ხედვებში ეს ორი სისტემა პრინციპულად განსხვავდება ერთმანეთისგან.

რომელიც ქართულ მათემატიკურ ანუ აზრით ენაში არის [გაუცვალა(Nმა,Nს,No, Nში)=t] საოპერაციო გამონათქვამი; 2. *ქალმა*, რომელიც ქართულ მათემატიკურ ანუ აზრით ენაში (ქალმა)_{Nმა} ტიპის მარტივი ჩასმის ოპერატორია; 3. *ქალს*, რომელიც ქართულ მათემატიკურ ანუ აზრით ენაში (ქალს)_{Nს} ტიპის მარტივ ჩასმის ოპერატორია; 4. *მსხალი*, რომელიც ქართულ მათემატიკურ ანუ აზრით ენაში (მსხალი)_{No} ტიპის მარტივი ჩასმის ოპერატორია; 5. *მსხალში*, რომელიც ქართულ მათემატიკურ ენაში (მსხალში)_{Nში} ტიპის მარტივი ჩასმის ოპერატორია. - ახლა ვნახოთ, როგორ ოპერირებენ ეს ოპერატორები:

(1) *ქალმა გაუცვალა* =

$$= (\text{ქალმა})_{Nმა} + {}^4[\text{გაუცვალა}]_{(Nმა, Nს, No, Nში)=t} - {}^3[\text{ქალმა გაუცვალა}]_{(Nს, No, Nში)=t}$$

ფორმულიდან ჩანს, რომ *ქალმა*, როგორც Nმა ტიპის ჩასმის ოპერატორი ოპერირებს *გაუცვალა* სიტყვით მოცემულ 4-ადგილიან წინადადებაზე, რომლის საოპერაციო ადგილთა ტიპებია - Nმა, Nს, No, Nში და შედეგად გვაძლევს *ქალმა გაუცვალა* წინადადებას, რომელიც ახალა უკვე არის 3-ადგილიანი წინადადებაა. ანუ, *ქალმა* ოპერირებით *გაუცვალა* წინადადების Nმა ადგილი ჩაიკეტა და წარმოიქმნა 3-ადგილიანი *ქალმა გაუცვალა* წინადადება, რომლის საოპერაციო ადგილთა ტიპებია Nს, No, და Nში.¹¹

(2) *ქალს გაუცვალა* =

$$= (\text{ქალს})_{Nს} + {}^4[\text{გაუცვალა}]_{(Nმა, Nს, No, Nში)=t} = {}^3[\text{ქალს გაუცვალა}]_{(Nმა, No, Nში)=t}$$

(3) *მსხალი გაუცვალა* =

$$= (\text{მსხალი})_{No} + {}^4[\text{გაუცვალა}]_{(Nმა, Nს, No, Nში)=t} = {}^3[\text{მსხალი გაუცვალა}]_{(Nმა, Nს, Nში)=t}$$

(4) *მსხალში გაუცვალა* =

$$= (\text{მსხალში})_{Nში} + {}^4[\text{გაუცვალა}]_{(Nმა, Nს, No, Nში)=t} - {}^3[\text{მსხალში გაუცვალა}]_{(Nმა, Nს, No)=t}$$

¹¹ აქ აღწერილ ოპერირებას განაპირობებს ის, რომ, ერთი მხრივ, *კაცმა* Nმა ტიპის ოპერატორია, მეორე მხრივ კი ის, რომ *გაუცვალა* წინადადების ერთ-ერთი საოპერაციო ადგილი Nმა ტიპისაა.

ამგვარად, კ.ფხაკაძის მიხედვით, ნებისმიერი $N\alpha$ ოპერატორის ($N\alpha \in \{N_0, N_{01}, N_{01კენ}, N_{01გან}, N_{01თვის}, N_{01თ}, N_{01დან}, N_1, N_{ში}, N_{ზე}, N_{ად}, N_{ამდე}, N_{თან}\}$) განსაზღვრის არეა იმ და მხოლოდ იმ ზმნური ფრაზების ანუ, ჩვენ შემთხვევაში, წინადადებების სიმრავლე, რომელთაც ერთი მაინც $N\alpha$ ტიპის ღია ადგილი აქვთ. ასეთ წინადადებებს ჩვენ ზოგჯერ $[S]_{(N\alpha)}$, ზოგჯერ კი $- [V]_{(N\alpha)}$ გამოსახულებით აღვნიშნავთ. ამასთან, $^k[S]_{(N\alpha)}$ -ით, ზოგჯერ კი $- ^k[V]_{(N\alpha)}$ გამოსახულებით აღვნიშნავთ ასეთ k -ადგილიან ($k \in \{1, 2, \dots\}$) წინადადებებს. ამგვარად, ახლა, მზად ვართ უკვე N ტიპის ოპერატორების ოპერირების შემდეგი ზოგადი ფორმულის ჩასაწერადაც და გასაგებადაც:

$$N\alpha + ^k[V]_{(N\alpha)} = ^k[V]_{(N\alpha)} + N\alpha = N^{-k} [N\alpha + V]_{(-)}$$

უკვე ვნახეთ, რომ კ.ფხაკაძის ლოგიკური გრამატიკაში ქართული არსებითი სახელების ბრუნვითი ფორმებით, როგორც შემამოკლებელი სიმბოლოებით, ქართული ენაში რეალიზებულია ჩასმის ოპერატორები. თუმცა N ოპერატორების შინაარსის უფრო სრულად გასახსნელად ჩვენ გვჭირდება აგრეთვე არსებითი სახელის სიმრავლურ-სემანტიკური გააზრება. ძალიან მოკლედ ამის შესახებ: პირველად [49]-ში კ.ფხაკაძის მიერ გამოითქვა მოსაზრება იმის თაობაზე, რომ ქართული არსებითი სახელებით აზროვნებაში მოგვეცემა ამ სახელებითვე ძირეულად განსაზღვრული სიმრავლეები და რომ ეს არსებითი სახელები წინადადებებში გაიგებიან როგორც ამ სიმრავლეებზე განსაზღვრული ან ცვლადი, ან კონსტანტა, ან თავად ეს სიმრავლე. შემდგომში ეს მოსაზრება კიდევ უფრო დასაბუთდა და გამყარდა, რითაც კ.ფხაკაძის მათემატიკური მიდგომები ძირეულად აცდა მონტეგიუს მიდგომებს. ასე მაგალითად, ქართულ მათემატიკურ ენაში, არსებითი სახელებით *‘ადამიანი’*, *‘მსხალი’* მოგვეცემა ყველა ადამიანების, შესაბამისად, ყველა მსხლების სიმრავლე. ამ სიმრავლეს, ჩვენ, უბრალო შეთანხმების ძალით, *{ადამიანი}*, *{მსხალი}* ჩანაწერებით აღვნიშნავთ. ამასთან, ჩვენთვის, *{ადამიანი}* სიმრავლე, ისევე როგორც *‘ადამიანი’* სიტყვით ქართულში მოცემული ცნება, არამირითადი ანუ წარმოებული ანუ განსაზღვრებადი სიმრავლეა, მაშინ როდესაც *{მსხალი}* სიმრავლე, ისევე

როგორც 'მსხალი' სიტყვით ქართულში მოცემული ცნება, ძირითადი ანუ პირველსაწყისი ანუ არაგანსაზღვრებადი სიმრავლეა. ანუ, ქართულში, ისევე როგორც ქართულ აზრით ანუ მათემატიკურ ენაში სიმრავლე *{ადამიანი}* შეიძლება განვსაზღვროთ *{კაცი, ქალი, ზავში}* სიმრავლეების გაერთიანებით. მართლაც, მაღალი ალბათობით, შეიძლება შევთანხმდეთ, რომ *{ადამიანი}={კაცი}∪{ქალი}∪{ზავში}* ჭეშმარიტი სიმრავლური ტოლობაა. თუმცა, იმას, რომ ქართულ ენაში *{მსხალი}* სიმრავლე შეუძლებელია რაიმე სიტყვებით მოცემული სიმრავლეებით განისაზღვროს, დაზუსტებით ვამბობთ, რადგან კ.ფხაკაძის მიხედვით 'მსხალი' სიტყვით ქართულში არაგანსაზღვრებადი ანუ ძირეული ანუ საწყისი სიმრავლური მოცემულობა გვეძლევა.¹²

ახლა ვნახოთ, რა როლს თამაშობენ ეს სიმრავლური მიდგომები N ოპერატორებში ანუ, ვნახოთ, კ.ფხაკაძის სხვადასხვა ლოგიკური ბრუნვითი ფორმები რა შინაარსულ დატვირთვებს იღებენ თავის თავზე მარტივი ზმნით აგებულ ქართულ მარტივ წინადადებებში. ანუ, როგორც უკვე აღვნიშნეთ, კ.ფხაკაძის მიხედვით, ქართული არსებითი სახელებით აზროვნებაში მოგვეცემა ამ სახელითვე განსაზღვრული სიმრავლე და, ამასთან, ეს არსებითი სახელი წინადადებებში გაიგება ან როგორც თავად ეს სიმრავლე, ანდა როგორც ამავე სიმრავლეზე განსაზღვრული ან ცვლადი, ან კონსტანტა. აღნიშნულის დასასაბუთებლად მოვიყვანოთ რამოდენიმე მაგალითი:

ადამიანი მოკვდავია - ყველა ადამიანი მოკვდავია (1)

ადამიანი მოკვდა - ყველა ადამიანი მოკვდა (2)

ქართული ენის ცოდნა საკმარისია იმ დასკვნის გასაკეთებლად, რომ (1) ხაზში განთავსებულ წინადადებებს ერთი და იგივე ფაქტობრივი შინაარსი

¹² ერთ-ერთი მიზეზი იმისა, რომ *{მსხალი}* ჩანაწერით ჩვენ აღვნიშნეთ ყველა მსხლების სიმრავლე და არა 'მსხალი' სიტყვის შემცველი 1 ელემენტის სიმრავლე, სწორედ ისაა, რომ იგი, როგორც ძირითადი სიმრავლე არ შეიძლება იქნეს მოცემული არც მისი ელემენტების ჩამოთვლის გზით, არც მანამდე არსებული მისი განმსაზღვრელი თვისების მინიშნებით და არც სხვა რაიმე სახის რეკურსიული განსაზღვრებით.

აქვთ, ხოლო (2) ხაზში განთავსებული წინადადებების შინაარსი არსებითად განსხვავდება ერთმანეთისაგან. კერძოდ, (1) ხაზის წინადადებები, მიუხედავად იმისა, რომ მათგან მეორე პირველის განზოგადებული ფორმაა, გვატყობინებენ, რომ მოკვდავობის თვისება ყველა ადამიანს მიემართება, მაშინ როდესაც, მეორე ხაზის პირველი წინადადება გვატყობინებს ერთი რომელიღაც ადამიანის მოკვდინებაზე, ხოლო მეორე, რომელიც პირველის განზოგადებული ფორმაა, გვატყობინებს რომ მოკვდინების ფაქტი შეეხება არა მხოლოდ ერთ რომელიღაც ადამიანს, არამედ ნებისმიერ ანუ ყველა ადამიანს. - ბუნებრივია ისმის კითხვა - რა არის იმის მიზეზი, რომ (1) ხაზში მოცემულ წინადადებაში ყველა სიტყვა ადამიანს არ ანზოგადებს, მაშინ როდესაც, (2) ხაზში ანზოგადებს? - საქმე ისაა, რომ 'ადამიანი მოკვდავია' წინადადებაში 'ადამიანი' გაიგება როგორც {ადამიანი} სიმრავლეზე განსაზღვრული ცვლადი. ამ სიტყვის ასეთ გაგებას (ადამიანი^v) -თი აღვნიშნავთ. ანუ, ამ მიდგომით 'ადამიანი მოკვდავია' გაიგება, როგორც ეს ქვემოთაა მოცემული:

$$(ადამიანი^v)_{No} + [მოკვდავია(No)=t] = [მოკვდავია(ადამიანი^v)=t] = x \in \{ადამიანი\}; მოკვდავია(x)=t$$

ანუ, ამ მიდგომით, 'ადამიანი მოკვდავია' წინადადება არის მეტყველი პირის მიერ ჭეშმარიტ მოცემულობად მოწოდებული ცვლადზე დამოკიდებული გამონათქვამი, რაც, ანუ ცვლადის შემცველი გამონათქვამის ჭეშმარიტ მოცემულობად მოწოდება, ამ ცვლადს ანზოგადებს ანუ წინადადებას გვააზრებინებს როგორც ჭეშმარიტს ამ ცვლადის ნებისმიერი მნიშვნელობისათვის. - ეს არის მიზეზი იმისა, რომ (1) ხაზის ორივე წინადადებას ერთი და იგივე შინაარსი აქვთ. (2) ხაზის პირველ წინადადებაში კი 'ადამიანი' გაიგება როგორც {ადამიანი} სიმრავლეზე განსაზღვრული არასაკუთრივი კონსტანტა. ამ სიტყვის ასეთ გაგებას (ადამიანი^c) -თი აღვნიშნავთ. ანუ, ამ მიდგომით 'ადამიანი მოკვდა' გაიგება, როგორც ეს ქვემოთაა მოცემული:

$$(ადამიანი^c)_{No} + [მოკვდა(No)=t] = [მოკვდა(ადამიანი^c)=t] = a \in \{ადამიანი\}; მოკვდა(a)=t$$

ანუ, ამ მიდგომით, 'ადამიანი მოკვდა' წინადადება არის მეტყველი პირის მიერ ჭეშმარიტ მოცემულობად მოწოდებული კონსტანტაზე დამოკიდებული გამონათქვამი, რაც ამ წინადადებას გვააზრებინებს როგორც კერძო ანუ ამ კონსტანტის ერთ რომელიმე კერძო მნიშვნელობაზე ჭეშმარიტ გამონათქვამს. აღნიშნულის გათვალისწინებით 'ყველა ადამიანი მოკვდა' შემდეგნაირად აზრდება:

$$\forall x([ადამიანი(x)] \supset [მოკვდა(x)]) = [მოკვდა(\forall x[ადამიანი(x)])] = \forall a \in [ადამიანი]: [მოკვდა(a)]$$

ანუ, ამ მიდგომით, 'ყველა ადამიანი მოკვდა' გაგებულია, როგორც a კონსტანტით {ადამიანი} სიმრავლეზე განზოგადებული [მოკვდა(a)] გამონათქვამი, რაც ცხადია, 'მოკვდა' მიმართებაში აყენებს ყველა ანუ ნებისმიერ 'ადამიანს'.

ამგვარად, ზემოთ წარმოდგენილი მსჯელობით დასაბუთდა, რომ (1) ხაზის წინადადებებში 'ადამიანი' გაიგება როგორც {ადამიანი} სიმრავლეზე განსაზღვრული ცვლადი, ხოლო (2) ხაზის წინადადებებში იგივე 'ადამიანი' გაიგება როგორც {ადამიანი} სიმრავლეზე განსაზღვრული კონსტანტა.

სახელების ასეთ გაგებასთან ერთად, ქართულში მარტივი ზედსართავები ანუ მარტივი A ტიპის სიტყვები, მსგავსად N ტიპის სიტყვებისა, სიმრავლეებად გაიგება. ასე მაგალითად, 'ყვითელი' გაიგება ყველა ყვითელი საგნების სიმრავლედ და მას {ყვითელი} ჩანაწერით აღვნიშნავთ. ზედსართავებით მოცემული სიმრავლეები განსხვავდება არსებითი სახელებით მოცემული სიმრავლეებისაგან იმით, რომ ისინი თვისების ხარისხის ზრდადობის მიმართ დალაგებული სიმრავლეებია. მართლაც 'ძალიან ყვითელი', აზრობრივად მისაღები და შინაარსულად გამართული ენობრივი გამოსახულებაა, მაშინ როდესაც 'ძალიან მსხალი' აზრობრივად გაუმართავი შინაარსია. ამის მიზეზი ისაა, რომ {ყვითელი}, როგორც დალაგებული სიმრავლე, 'ძალიან' ოპერატორს იღებს, {მსხალი} კი, როგორც დაულაგებელი სიმრავლე, - არა. ასევე, ლოგიკური გრამატიკით დადგინდა მარტივი ზედსართავების რთული მათემატიკური სტრუქტურა. კერძოდ, დადგინდა, რომ ნებისმიერი მარტივი A ზედსართავი აგებულია სიმრავლური თანაკვეთის 2 ადგილია ოპერაციით - (\cap) და {A} სიმრავლით.

კერძოდ, დადგინდა, რომ ნებისმიერი ასეთი A ზედსართავით მოკლდება ($\{A\}_r$) სახის რთული ენობრივ-მათემატიკური გამოსახულება, რომელიც გამომდინარე მისი სტრუქტურიდან მარჯვნივ 1-ადგილიანი ოპერატორია. უფრო გასაგები რომ იყოს: ქართული ენის ლოგიკური გრამატიკით 'ყვითელი' მარჯვნივ 1-ადგილიანი ოპერაციაა და იგი ამოკლებს $\{ყვითელი\}_r$ ფორმას, რომელიც აგებულია $\{ყვითელი\}$ -ით აღნიშნული სიმრავლით და r თანაკვეთის მათემატიკური ლოგიკური ოპერაციით. მართლაც, კ.ფხაკაძის სიმრავლურ-თორიული მიდგომებით ზედსართავული სახელური ფრაზა 'ყვითელი მსხალი' გაიგება როგორც $\{ყვითელი\}_r\{მსხალი\}$, ანუ როგორც ყვითელი საგნებისა და მსხლების სიმრავლეების თანაკვეთა, რაც ცხადია შესაბამისობაშია ქართულში არსებულ ვითარებასთან.

ლოგიკური გრამატიკის ჩვენთვის მნიშვნელოვანი საკითხების სპექტრალური მიმოხილვის მიზნით ძალიან მოკლედ შემამოკლებელი მორფემების შესახებ: დღეს არსებული ხედვით მორფემა არის ან დერივაციული ან ფლექსიური. დერივაციული მორფემა მოქმედებს რა ძირზე იძლევა ან იმავე მეტყველების ნაწილის ახალ მნიშვნელობას, ან სხვა მეტყველების ნაწილის ახალ სიტყვას. ფლექსიური მორფემა მოქმედებს რა სიტყვაზე უცვლის მას გრამატიკულ მნიშვნელობას. ლოგიკურ გრამატიკით კი დგინდება, რომ მორფემათა ეს კლასიფიცირება არასრულია. მაგ., მორფემა ((-)ს), სიტყვაში ((ყვითელ)ს) არ არის ფლექსიური, რადგან 'ყვითელს' არ არის 'ყვითელი' სიტყვით მოცემული ზედსართავის ბრუნვითი ფორმა. ამასთან, იგი არც დერივაციულია, რადგან იგი არც სხვა რომელიმე მეტყველების ნაწილია. მართლაც, ქართულში, 'ყვითელს' ამოკლებს 'ყვითელ Ns', რაც, ცხადია, ვერცერთი მეტყველების ნაწილის კუთვნილებად ვერ შეფასდება. ამგვარად, ქართული ენის ლოგიკური გრამატიკით მტკიცდება, რომ ((-)ს) მორფემა სიტყვაში 'ყვითელს' არც ფლექსიურია და არც დერივაციული და მას შემამოკლებელ მორფემას ვუწოდებთ. ამასთან, 'ყვითელს' ჩვენთვის შიდაენობრივი შემამოკლებელი სიტყვაა, რადგან მისით შემოკლებული

გამოსახლება სრულად იშლება თავად ენაშივე. - ასეთი სიტყვები ქართულში მრავლადაა. ასე მაგალითად, ასეთი შიდაენობრივი შემამოკლებელი სიტყვის მაგალითია აგრეთვე ზმნა 'ყვითელი', რომელიც ამოკლებს ენობრივ გამოსახულებას 'არის ყვითელი'.

ახლა ძალიან მოკლედ ზედსართავი (განმაზოგადებელი) სახელების ანუ A (Q) ტიპის სიტყვების ბრუნებისა და ოპერატორული ბუნების შესახებ: ლოგიკური გრამატიკის მიხედვით, ქართულში, ზედსართავ (განმაზოგადებელ) ანუ A (Q) სახელს აქვს 3 ბრუნვა. ესენია: A₀(Q₀), A₋(Q₋), A_{მა}(Q_{მა}) ბრუნვები. არაბრუნებადი A (Q) სახელები აღნიშნება A*-თი (Q*-თი). ამასთან, დადგინდა, რომ A (Q) სახელები მარცხნივ ერთდაგილიანი ოპერატორებია, რომლებიც უშუალოდ N ტიპის სიტყვებზე მოქმედებენ. დადგინდა ისიც, A (Q) სახელების ბრუნვის ნიშნებით მათი განსაზღვრის არეები მოგვეცემა. ასე მაგალითად: 1. A₀ და Q₀ ოპერატორების განსაზღვრის არეა {N₀, N_{ის}, N_{ისკენ}, N_{ისგან}, N_{ისთვის}, N_{ით}, N_{იდან}}; 2. A₋ და Q₋ ოპერატორების განსაზღვრის არეა {N_ს, N_{ში}, N_{ზე}, N_{ად}, N_{ამდე}, N_{თან}}; 3. A_{მა} და Q_{მა} ოპერატორების განსაზღვრის არეა {N_{მა}}. ასევე, დადგინდა ისიც, რომ A₀, A₋A_{მა}, A* (Q₀, Q₋, Q_{მა}, Q*) მოქმედებენ რა განსაზღვრის არეში შემავალ და მათ მარჯვნივ განთავსებულ N_α-ტიპის რომელიმე AN_α-ფრაზაზე შედეგად იძლევიან N_α-ტიპის AN_α-ფრაზას (QN_α-ფრაზას), რაც უზრუნველყოფს ამ ოპერატორების მიმართ {N₀, N_{ის}, N_{ისკენ}, N_{ისგან}, N_{ისთვის}, N_{ით}, N_{იდან}, N_ს, N_{ში}, N_{ზე}, N_{ად}, N_{ამდე}, N_{თან}, N_{მა}} ტიპთა სიმრავლის ჩაკეტილობას.

ამგვარად, ახლა უკვე მზად ვართ აღვწეროთ ქართული ენის I საფეხურის მათემატიკური თეორია - ქართული ენის ძირეული ნაწილის ლოგიკური გრამატიკა, რომელიც ჩვენი პროგრამისტული გადაწყვეტებისთვის სასაფუძვლოდ მნიშვნელოვანი თეორიაა:

1. საგნობრივი ასოები (ცვლადები და კონსტანტები) ანუ N ტიპის სიტყვები: N₀, N_{ის}, N_{ისკენ}, N_{ისგან}, N_{ისთვის}, N_{ით}, N_{იდან}, N_ს, N_{ში}, N_{ზე}, N_{ად}, N_{ამდე}, N_{თან}, N_{მა};

2. ფუნქციონალური კონსტანტები ანუ A ტიპის სიტყვები: A-, A₀, A_{მა}, A*;
3. კვანტორული ნიშნები ანუ Q ტიპის სიტყვები: Q-, Q₀, Q_{მა}, Q*;
4. საოპერაციო კონსტანტები ანუ V ტიპის სიტყვები: [V_m(N₁,...,N_m)=t]
(N_k∈{N₀, N_{ის}, N_{ისკენ}, N_{ისგან}, N_{ისთვის}, N_{ით}, N_{იდან}, N_ს, N_{ში}, N_{ზე}, N_{ად}, N_{ამდე}, N_{თან}, N_{მა}}, k∈{1, ...,m}).

სახელური ფრაზების ამგები წესი:

1. ნებისმიერი N_a ტიპის სიტყვა არის N_a ტიპის სახედსართავო ანუ ღია სახელური ფრაზა;

2. A_b + N_a არის N_a ტიპის ღია სახელური ფრაზა მაშინ, და მხოლოდ მაშინ, როცა N_a ტიპი არის A_b ოპერატორის განსაზღვრის არის ელემენტი და N_a ტიპის ღია სახელური ფრაზა;

3. Q_c + N_a არის N_a ტიპის კვანტორული ანუ ჩაკეტილი სახელური ფრაზა მაშინ, და მხოლოდ მაშინ, როცა N_a ტიპი არის Q_c კვანტორული ნიშნის განსაზღვრის არის ელემენტი და N_a ტიპის ღია სახელური ფრაზა.

ზმნური ფრაზების (წინადადებების) ამგები წესი:

1. [V_m(N₁,...,N_m)=t]^m პრედიკატული კონსტანტა არის m ადგილიანი მარტივი დაუსრულებელი ზმნური ფრაზა.

2. [V_m(N₁,...,N_m)=t]^m + N_a არის m – 1 ადგილიანი ზმნური ფრაზა (ან მარტივი, ან შედგენილი), მმ, როცა N_a ∈{N₁,...,N_m}. ამასთან, იგი არის: 1. თუ m – 1 = 0, დასრულებული ზმნური ფრაზა ანუ დასრულებული წინადადება;

2. თუ m – 1 > 0, დაუსრულებელი ზმნური ფრაზა ანუ დაუსრულებელი წინადადება;

3. N_a + [V_m(N₁,...,N_m)=t]^m არის იმავე ტიპის ზმნური ფრაზა, რა ტიპის ზმნური ფრაზაცაა [[V_m(N₁,...,N_m)=t]^m + N_a

ქართული ენის ძირეული ნაწილის ზემოაღწერილი მათემატიკური თეორია ანუ ქართული ენის ძირეული ნაწილის ლოგიკურ გრამატიკა ჩვენი პროგრამისტული გადაწყვეტებისათვის სასაფუძვლოდ მნიშვნელოვანია და სწორედ ამიტომ მის მთელ რიგ შემადგენელ საკითხებზე ზემოთ საკმაოდ ვრცელი საუბრები გვქონდა.

თავი 2. შედეგები და მათი განსჯა

2.1. აფხაზური ენის თვითგანვითარებადი კორპუსი და მასში ჩადგმული აფხაზური ვებ კლავიატურა

საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების საწავლო-სამეცნიერო ცენტრის თვითგანვითარებადი კორპუსის უკვე შემუშავებულ საცდელ-სამომხმარებლო ვერსიას ანუ ქართული ჭკვიანი ვებ-კორპუსის საცდელ-სამომხმარებლო ვერსიას (იხილეთ www.corpus.ge მისამართზე) ქართული, ინგლისური, და რუსული ენის კორპუსებთან ენებთან ერთად სადოქტორო კვლევის მიმდინარე ეტაპზე საცდელ-სამომხმარებლო სახით დაემატა აგრეთვე აფხაზური ენის კორპუსიც. ამგვარად, ახლა უკვე შესაძლებელია ჩვენს უნივერსალურ კორპუსში მოიძებნოს მასალა აფხაზურ ენაზეც. მიება შესაძლებელია როგორც სიტყვის, ასევე ფრაზის მეშვეობითაც [11].

თვითგანვითარებად ქართულ ჭკვიან ვებ-კორპუსში ცალკე კორპუსად ჩადგმული აფხაზური ენის კორპუსის განსავითარებლად ცხადია საჭიროა აფხაზურენოვანი წყაროები, რაც პერმანენტულად და, ასევე, ავტომატურად ამ წყაროებიდან აფხაზურენობრივი ტექსტების კორპუსში შემოტანის საშუალებას იძლევა. კერძოდ, გარკვეული პერიოდულობით ანუ დროის წინასწარ ფიქსირებული ინტერვალებით ხდება საინტერნეტო წყაროს შემოწმება და, შესაბამისად, იმავე პერიოდულობით ირკვევა არის თუ არა წყაროში ბოლო შემოწმების შემდგომი რაიმე ახალი მასალა. ამასთან, თუ სისტემამ დააფიქსირა რომ წყარო მოიცავს ახალ მასალას, მყისიერად ხდება ამ მასალის წამოღება ანუ ჩვენი კორპუსის მონაცემთა ბაზის სტრუქტურაზე მისი მორგება, დამუშავება, და შენახვა. ახალ მონაცემებს ვინახავთ ტექსტის სახით. შემდგომ [11]:

1. თითოეული ასეთ ტექსტს ვშლით წინადადებებად და ვინახავთ მათ წინადადებათა ბაზაში ცალკე წინადადებად. ამასთან, თითოეულ

წინადადებას უფიქსირდება ანუ უწყვილდება მისი დედა ტექსტი ანუ ტექსტი, რომლიდანაც ეს წინადადებაა წამოღებული;

2. თითოეულ ასეთ დედა ტექსტთან უკვე დაწყვილებულ წინადადებას ვშლით სიტყვებად და ვინახავთ მათ სიტყვათა ბაზაში ცალკე სიტყვებად. ამასთან, თითოეულ სიტყვას უფიქსირდება ანუ უწყვილდება მისი დედა წინადადება ანუ წინადადება, რომლიდანაც ეს სიტყვაა წამოღებული.¹³

გარდა ამისა [11]:

1. თითოეულ მონაცემის შენახვისას ფიქსირდება მისი მოძიების თარიღი ანუ მისი დედა ტექსტის ბაზაში შემოტანის თარიღი. ანუ, ბაზაში შემოსული ტექსტის წინადადებად და სიტყვებად ჩაშლისას ყველა ამ მონაცემს მიჰყვება დედა ტექსტის კორპუსში შემოსვლის თარიღი;

2. თითოეულ მონაცემის შენახვისას მონაცემი დარდება კორპუსში ანუ ბაზაში უკვე არსებულ იმავე ტიპის უნიკალურ მონაცემებს და: ა) თუ ეს მონაცემი ამ ტიპის ახალი უნიკალური მონაცემია, საბაზო მონაცემთა ეს ტიპი ფართოვდება ამ ახალი უნიკალური მონაცემით, რომლის შემოსვლათა სიხშირის მაჩვენებლად ცხადდება ერთი; ბ) ხოლო, თუ ეს მონაცემი ამ ტიპის ახალი უნიკალური მონაცემი არ არის, საბაზო მონაცემთა ამ ტიპის

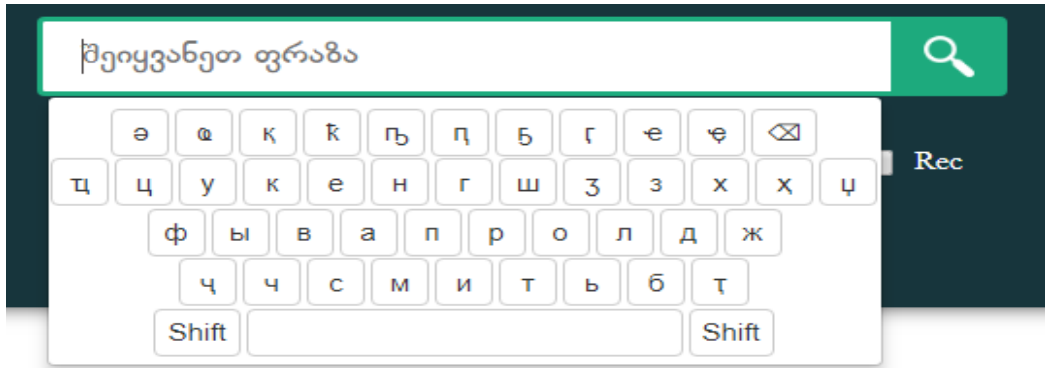
¹³ ქართულში ტექსტის ჩაშლა გარდა ზემოწარმოდგენილი 1-2 ბიჯებისა ხდება აგრეთვე ქვემოთ წარმოდგენილი მე-2 და მე-3 ბიჯებით: 2. თითოეულ ასეთ დედა სიტყვასთან უკვე დაწყვილებულ სიტყვას ვშლით მარცვლებად და ვინახავთ მათ მარცვალთა ბაზაში ცალკე მარცვლებად. ამასთან, თითოეულ მარცვალს უფიქსირდება ანუ უწყვილდება მისი დედა სიტყვა ანუ სიტყვა, რომლიდანაც ეს მარცვალია წამოღებული; 3. თითოეულ ასეთ დედა სიტყვასთან უკვე დაწყვილებულ მარცვალს ვშლით მარცვლოიდებად ანუ მისი შემადგენელ ასოების დაინდექსირებულ ცალეებად და ვინახავთ მათ მარცვლოიდების (დაინდექსირებული ასოების) ბაზაში ცალკე დაინდექსირებულ ასოდ ანუ მარცვლოიდად. ამასთან, თითოეულ ასეთ მარცვლოიდს (დაინდექსირებულ ასოს) უფიქსირდება ანუ უწყვილდება მისი დედა მარცვალი ანუ მარცვალი, რომლიდანაც ეს მარცვალია წამოღებული და, აგრეთვე, ამ დაინდექსირებული ასოს უინდექსო ანუ ანბანური ვერსია (მხედველობაში გვაქვს წერითი ენის ანბანი). - აფხაზურში, ჯერ-ჯერობით, ეს აქ მე-2 და მე-3 ბიჯით წარმოდგენილი ჩაშლები არ გვაქვს.

უნიკალურ მონაცემთა სიაში ამ ახლად შემოსული მონაცემის სიხშირის მაჩვენებელი იზრდება ერთით. ამგვარად, ეს ყველაფერი, საშუალებას იძლევა თითოეული ახალი მონაცემისთვის დაფიქსირდეს როგორც მისი შემოსვლის თარიღი, ისე სიხშირე, ასევე წყარო და მონაცემის ენობრივი კუთვნილება, რაც აფხაზური მონაცემებისათვის მათზე მიბმული აფხაზური ენის ინდექსის მეშვეობით ხდება [11].

დღეისათვის მოძიებულია და კორპუსზე ავტომატურ რეჟიმში უკვე მიერთებული შემდეგი აფხაზურენოვანი წყაროები: www.apsnypress.info; <http://abkhazia.gov.ge>; http://gov.ge/index.php?lang_id=ABK. ამასთან, ამ წყაროებიდან სულ ჯამურად უკვე წამოღებულია 7650858 აფხაზური სიტყვა, რომელთაგან 817872 განსხვავებულია (მონაცემები აღებულია 2019 წლის 18 დეკემბრის ჩვენების მიხედვით). როგორც უკვე აღვნიშნეთ ეს მონაცემები ყოველდღიურად განიცდის მატებას. მატება დამოკიდებულია წყაროებში ახალი მასალის განთავსების ინტენსივობაზე [11].

ახლა ძალიან მოკლედ აფხაზურ ენის თვითგანვითარებად კორპუსში ჩადგმულ აფხაზურ ვებ კლავიატურაზე: ჩვენი ვებ-კორპუსის საიტზე (განთავსებულია www.corpus.ge მისამართზე) აფხაზურ ენაზე გადასვლისას საძიებო სისტემის ტექსტური ველი ილჭურვება აფხაზური კლავიატურით (იხ., ქვემოთ, სურათი 3). ეს მომხმარებელს საშუალებას აძლევს სისტემას მასზე ამ უკვე მიდგმული კლავიატურით მიაწოდოს მის მიერვე აკრეფილი ის საძიებო სიტყვა ან ფრაზა, რომლის მიხედვითაც მას აფხაზურ კორპუსში მისთვის სასურველი ძიების წამოწყება სურს.¹⁴ ამასთან, სასურველი საძიებო სიტყვის/ფრაზის ჩაწერის შემდგომ, ძეხნის დასაწყებად საჭიროა დავაწვეთ ძეხნის ღილაკს და კორპუსი დაგვიბრუნებს ჩვენს მიერ მითითებულ სიტყვაზე/ფრაზაზე მოძიებულ მასალას [11].

¹⁴ აფხაზური კლავიატურის შესაქმნელ წყაროდ გამოყენებულ იქნა ვიკიპედია. კერძოდ კი https://en.wikipedia.org/wiki/Abkhaz_alphabet მისამართზე განთავსებული მასალა.



სურათი 3 - აფხაზური ვებ კლავიატურა

სიტყვით წამოწყებული ძიების შემთხვევაში ამ სიტყვის შემცველი წინადადებებთან და მათზე მიბმულ წინადადების შემცველ წყაროს ბმულთან ერთად გვიბრუნდება სტატისტიკური ინფორმაცია თავად ამ საძიებო სიტყვაზე. კერძოდ, კორპუსში ჩადგმული სტატისტიკური მთვლელები და ინფორმანტები გვაწვდიან სიტყვასთან დაკავშირებულ მონაცემებს (იხ., ქვემოთ, სურათი 4). ეს მონაცემებია: 1. კორპუსში ამ სიტყვის შემოსვლათა რაოდენობა ანუ კორპუსში სიტყვის სიხშირე; 2. სიტყვის ფარდობითი სიხშირე მიმდინარე მომენტისათვის; 3. სიტყვის შემოსვლის თარიღი ანუ პირველად როდის დაფიქსირდა ეს სიტყვა მონაცემთა ბაზაში.

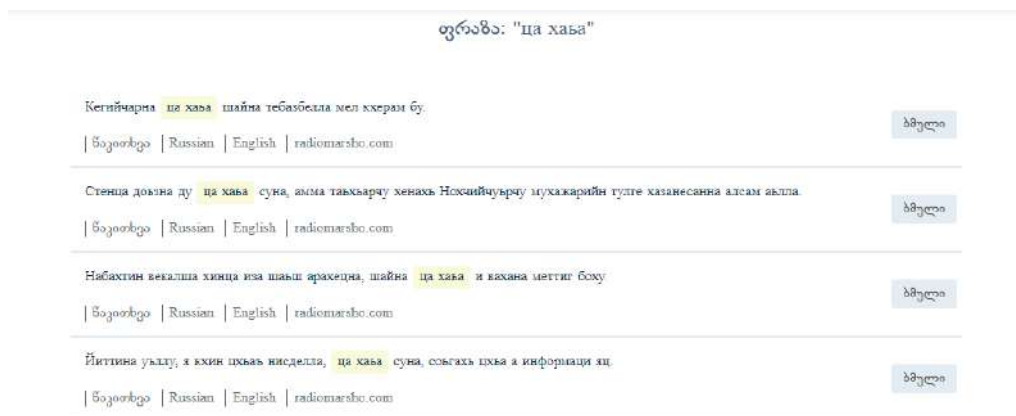


ფრაზა: "აღვსა" | რაოდენობა: 68 | ფარდობითი სიხშირე: 0.0000242175% | შემოსვლის დრო: 2017-04-15 15:38:13

სურათი 4 - აფხაზური კორპუსი მოქმედებაში

სურათ_4-ზე წარმოდგენილია აფხაზური სიტყვა „აღვსა“-თ წამოწყებული ძიების შედეგი, საიდანაც ჩანს რომ ამ სიტყვის სიხშირეა - 68;

ფარდობითი სიხშირეა - 0, 0000242175%; შემოსვლის თარიღია - 2017-04-15 ანუ 2017 წლის 15 აპრილი. გარდა ამისა, როგორც უკვე აღვნიშნეთ, გვიბრუნდება აგრეთვე საძიებო სიტყვის/ფრაზის შემცველი ყველა უნიკალური წინადადება, იმ წინადადებებიდან, რაც ძიების დროის მომენტისათვის კორპუსშია, მათი შემცველი წყაროების ბმულებთან დაწყვილებული სახით. (იხ., ქვემოთ, სურათი_5) [11].



სურათი 5 - აფხაზური კორპუსი ფრაზით ძიებისას

როგორც ზემოთ წარმოდგენილი სურათიდან ჩანს ფრაზით წარმოებულ ძებნა იძლევა: 1. ფრაზის შემცველ წინადადებებს პირდაპირი ტექსტობრივი სახით; 2. საინტერნეტო ბმულების სახით იმ წყაროებს, საიდან არის დაძებნილი წინადადება წამოღებული.

გარდა ამისა, დაძებნილ ფრაზაზე მიდგმული წაკითხვის რეჟიმის ჩართვით ანუ „წაკითხვა“ ღილაკის გააქტიურებით კორპუსში ჩადგმული აფხაზური მკითხველი საშუალებას აძლევს მომხმარებელს მოისმინოს დაძებნილი წინადადების აფხაზურად გახმოვანებული ვერსია (უფრო დეტალურად იხილეთ ქვემოთ) [11].

2.2. აფხაზურ კორპუსში ჩადგმული პირველი და ჯერ-ჯერობით ერთადერთი აფხაზური მკითხველი სისტემის საცდელი ვერსია

როგორც უკვე აღვნიშნეთ, და აქ ჩვენ ვეყრდნობით დღეს ჩვენს ხელთ არსებულ მონაცემებს, კორპუსში აფხაზური ტექსტების კომპიუტერული

კითხვის უზრუნველყოფის მიზნით ჩვენ მიერ შემუშავებული აფხაზური მკითხველის საცდელი ვერსია, რომელიც ტექსტის წასაკითხად იყენებს ჩვენს მიერვე შემუშავებულ ქართულ მკითხველს, არის აფხაზური ენის პირველი და ჯერ-ჯერობით ერთადერთი კომპიუტერული მკითხველი სისტემა. ამგვარად, ჩვენი მიდგომით, აფხაზური ტექსტის წასაკითხად პირველ რიგში ხდება ამ აფხაზური ტექსტის გადაწერა მის ხმოვან ეკვივალენტ ქართულ ტექსტში. ამ მიზნით ჩვენ ვსარგებლობთ აფხაზური ასო-ბგერების ქართლი შესატყვისების ქვემოთ წარმოდგენილი ცხრილით (იხ., ქვემოთ, ცხრილი_4).¹⁵

აფხაზური ტექსტის ქართულად გადაწერის ალგორითმი ეყრდნობა ცხრილ_4-ს. კერძოდ, აფხაზურ მასალაში ნახულობს გადასაწერ აფხაზურ ანბანურ ასო ნიშანს (პრიორიტეტი ენიჭება ორ სიმბოლოიანებს როგორებიცაა гъ, гə, ҕъ, ҕə, дə, жъ, жə, зə, къ, кə, қъ, қə და ა.შ..) და მას ანაცვლებს მისი ცხრილური ქართული მეწყვილეთი (ან ერთი ასოთი, ან ასოთა კომპლექსით) და ასე აგრძელებს მანამ სანამ ყველა აფხაზური სიმბოლო არ გადაიწერება მათი ქართული ცხრილური შესატყვისებით [11].

უფრო გასაგები რომ იყოს, მოვიყვანოთ აფხაზური სიტყვის ქართულად გადაწერის მაგალითი: დავუშვათ, გადასაწერად მოცემულია სიტყვა ахатҭа (цəа). ცხრილზე დამყარებული შესატყვისების დაძებნის ალგორითმული წესით, ამ სიტყვის შემადგენელი აფხაზური ანბანური ნიშნები ქართულად ქვემოთ მოცემული სქემის მიხედვით გადაიწერება:

a -> ა, x -> ხ, a -> ა, т -> წ, a -> ა

(цə -> ცუ, а -> ა)

ამგვარად, აფხაზური „ахатҭа“ („цəа“) სიტყვის ქართულად გადაწერის ანუ ტრანსლიტერაციის შედეგი მიიღება „ახაწა“ („ცუა“), რომელიც

¹⁵ ცხრილის შედგენაში კონსულტაციების სახით მნიშვნელოვანი დახმარება გაგვიწია აფხაზური ენის ცნობილმა ქართველმა სპეციალისტმა რიტა მარღანიამ, რომელსაც აფხაზური ენის, როგორც მშობლიური ენის, ცხადია სრულყოფილი ცოდნა აქვს.

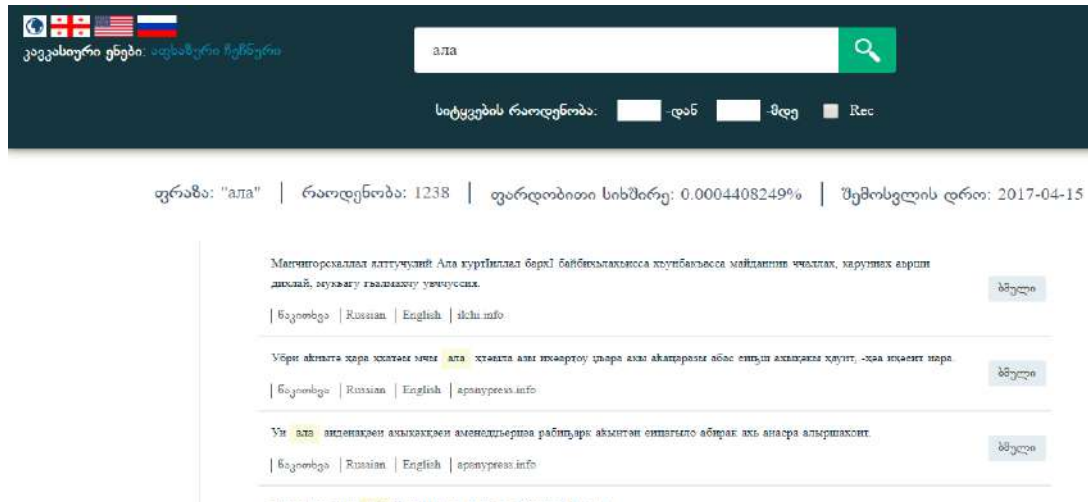
გადასაწერი აფხაზური სიტყვის ქართული ხმოვანი შესატყვისი ანუ ექვივალენტია [11].

აფხაზური ასო-ბგერების ქართული შესატყვისების ცხრილი								
	აფხაზური	ქართული		აფხაზური	ქართული		აფხაზური	ქართული
1	а	ა	22	кә	კუ	42	ф	ფვუ
2	б	ბ	23	қ	ქ	43	х	ხ
3	в	ვ	24	къ	ქ	44	хь	ხ
4	г	გ	25	қә	ქუ	45	хә	ხუ
5	гь	გ	26	к	ყ	46	х	ჰ
6	гә	გუ	27	кь	ყ	47	хә	ჰუ
7	д	დ	28	кә	ყუ	48	ц	ც
8	дь	დ	29	л	ლ	49	цә	ცუ
9	дә	დუ	30	м	მ	50	ц	წ
10	д	დ	31	н	ნ	51	цә	წუ
11	дә	დუ	32	о	ო	52	ч	ჩ
12	е	ე	33	п	პ	53	е	ჩუ
13	ж	ქე	34	пь	ფ	54	ч	ჭ
14	жь	ქ	35	р	რ	55	чә	ჭუ
15	жә	ქუ	36	с	ს	56	ш	შ
16	з	ზ	37	т	ტ	57	шь	შ
17	з	ძ	38	тә	ტუ	58	шә	შუ
18	зә	ძუ	39	т	თ	59	ы	იე
19	и	ი	40	тә	თუ	60	ә	იუ
20	к	კ	41	у	უ	61	и	ჯუ
21	кь	კ				62	иь	ჯ

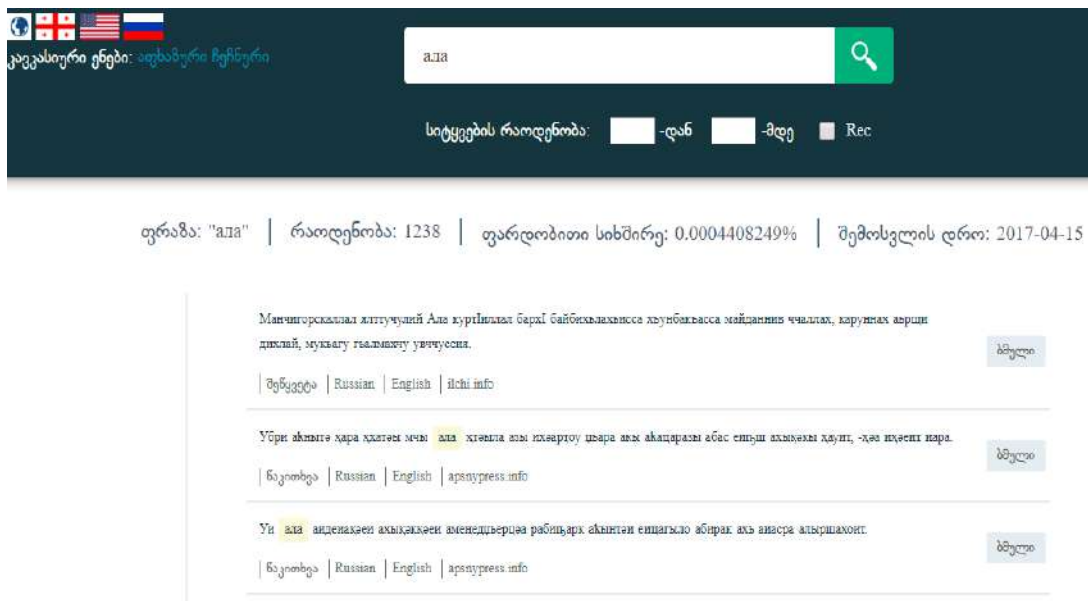
ცხრილი 4 - ქართულ აფხაზური ტრანსლიტერაციის ცხრილი

აფხაზური მკითხველი მოქმედებაში ანუ კითხვისას შეგიძლიათ იხილოთ www.corpus.ge მისამართზე (იხ., ქვემოთ, სურათი_6 და სურათი_7). ამისათვის: 1. კორპუსის მთავარ გვერდზე კავკასიური ენების ჩამონათვალში უნდა ავირჩიოთ აფხაზური ენა; 2. ამის შემდეგ საძიებო ტესტურ ველში უნდა ჩავწეროთ აფხაზური სიტყვა ან ფრაზა და გავუშვათ ძებნაზე; 3. მოძიებულ წინადადებებს ქვედა მარცხენა კუთხეში დართული აქვს ლილაკი „წაკითხვა“, რომლის გააქტიურებით კორპუსი იწყებს შესაბამისი წინადადების კითხვას აფხაზურად; 4. ამასთან, თავად ეს ლილაკი მისი გააქტიურების შემდეგ

ტრანსფორმირდება (ანუ გარდაიქმნება) „შეწყვეტა“ ღილაკად, რომლის გააქტიურებითაც მომხმარებელს შეუძლია სისტემას შეაწყვეტინოს მანამდე უკვე მიმდინარე კითხვის პროცესი (იხ., სურათი_7) [11].



სურათი 6 - აფხაზური მკითხველი მოქმედებაში (1)



სურათი 7 - აფხაზური მკითხველი მოქმედებაში (2)

2.3. ჩეჩნური, ლეკური და ყაბარდოული ენების თვითგანვითარებადი კორპუსების პირველი საცდელი ვერსიები და მათში ჩადგმული ჩეჩნური, ლეკური და ყაბარდოული ვებ-კლავიატურები

ქართულ ჭკვიანი თვითგანვითარებადი ვებ-კორპუსი სადოქტორო კვლევის მიმდინარე ეტაპზე აფხაზურის გარდა საცდელი სახით უკვე

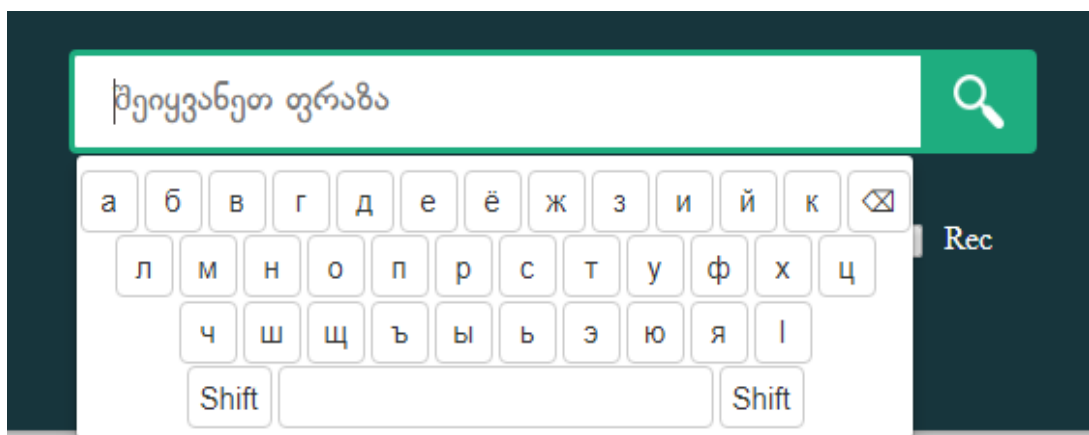
გაფართოვდა აგრეთვე სხვა კავკასიური ენებითაც. ესენია, ჩეჩნური, ლეკური და ყაზარდოული ენები. ეს გაფართოებები მოხდა იმ გამოცდილების გათვალისწინებით, რაც ჩვენ ამ მიმართულებით აფხაზურისთვის უკვე შევიძინეთ. ანუ, ეს გასაგებს ხდის, რომ ჩეჩნური, ლეკური, და ყაზარდოული ენების თვითგანვითარებადი კორპუსების ასაგებად გარკვეული დროის ინტერვალებში ხდება ამისთვის აუცილებელი ჩეჩნური, ლეკური, და ყაზარდოულ წყაროების შემოწმება და გარკვევა არის თუ არა წყაროში ბოლო შემოწმებიდან რაიმე ახალი მასალა. ამასთან, თუ სისტემამ დააფიქსირა რომ წყარო მოიცავს ახალ მასალას მყისიერად ხდება ამ მასალის წამოღება და ჩვენ მონაცემთა ბაზის სტრუქტურაზე მორგება მანამდე აფხაზურისთვის შემუშავებული შემუშავებული ტექსტების ავტომატურად დამამუშავებელი სისტემების მეშვეობით. ანუ, აქაც, მსგავსად აფხაზურისა, ახალ მონაცემებს ვინახავთ ტექსტის სახით; შემდგომ ვშლით ტექსტს წინადადებებად და ვინახავთ მათ ცალკე წინადადებად უნიკალურ წინადადებათა ბაზაში (ამასთან, თითოეული წინადადება წყვილდება მის დედა ტექსტთან ანუ იმ ტექსტთან, რომლიდანაც ის არის ამოღებული); წინადადებებს ვშლით სიტყვებად და ინახება სიტყვები ჩვენ მონაცემთა ბაზაში (ამასთან, თითოეული სიტყვა წყვილდება მის დედა წინადადებასთან ანუ იმ წინადადებასთან, რომლიდანაც ის არის ამოღებული. გარდა ამისა, ახალი სიტყვით ფართოვდება უნიკალური სიტყვათა ბაზა, ხოლო არაახალი სიტყვით ბაზაში ერთით იზრდება მისი სიხშირის მაჩვენებელი). გარდა ამისა, ყოველი მონაცემის შენახვისას ფიქსირდება მისი მოძიების თარიღი. მაგალითად ყოველი ახალი სიტყვის შენახვისას, ჩვენ ვინახავთ ამ ახალი სიტყვის კორპუსში პირველად შემოსვლის თარიღსაც, რაც, ცხადია, მეტად მნიშვნელოვანია გრძელვადიანი კვლევითი ინტერესებიდან გამომდინარე სიტყვის ქრონოგრაფიული და ისტორიოგრაფიული შესწავლის თვალსაზრისებით, რადგან აქ ჩვენ ამომწურავად გვიგროვდება მონაცემები სიტყვის პირველგამოჩენის დროზე, ნაშრომზე და ასევე მის პირველმთქმელზე ანუ ავტორზე [11]. ჯერ-ჯერობით:

1. ჩვენს მიერ ჩეჩნურ წყაროზე დაყრდნობით აგებული ჩეჩნური კორპუსი ჯამურად უკვე შეიცავს 121986 ჩეჩნურ სიტყვას, ამათგან განსხვავებულია 17798 სიტყვა (მონაცემები აღებულია 2019 წლის 18 დეკემბრის ჩვენების მიხედვით).

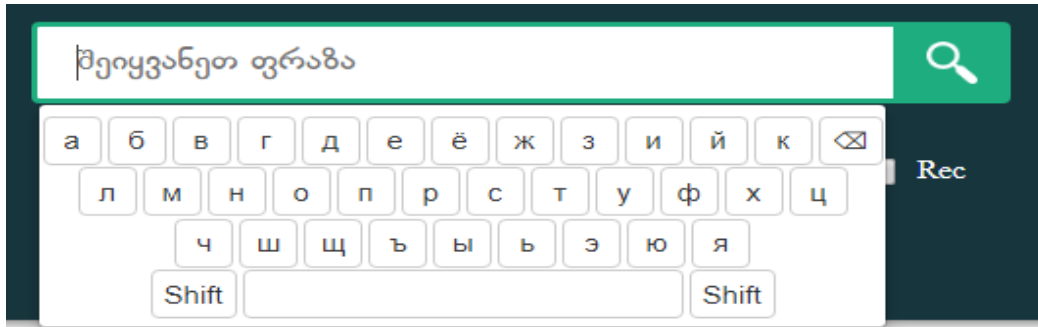
2. ჩვენს მიერ ლეკურ წყაროზე დაყრდნობით აგებული ლეკური კორპუსი ჯამურად უკვე შეიცავს 401448 ჩეჩნურ სიტყვას, ამათგან განსხვავებულია 130483 სიტყვა (მონაცემები აღებულია 2019 წლის 18 დეკემბრის ჩვენების მიხედვით).

3. ჩვენს მიერ ყაბარდოულ წყაროზე დაყრდნობით აგებული ყაბარდოული კორპუსი ჯამურად უკვე შეიცავს 81680 ჩეჩნურ სიტყვას, ამათგან განსხვავებულია 45036 სიტყვა (მონაცემები აღებულია 2019 წლის 18 დეკემბრის ჩვენების მიხედვით).

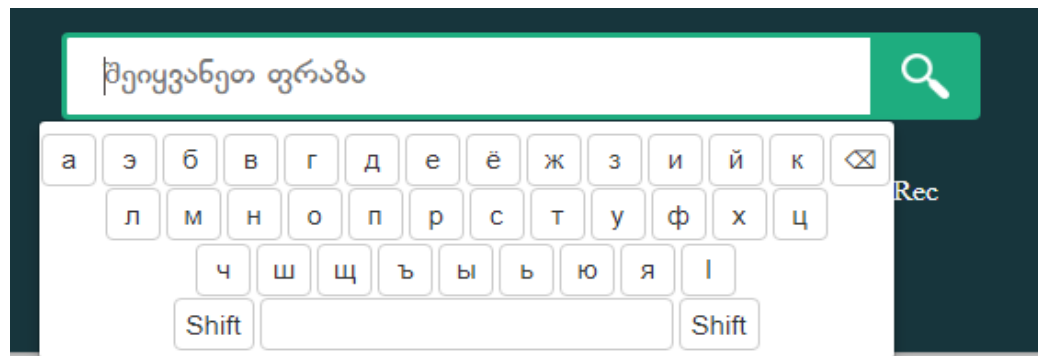
ხაზგასასმელია ისიც, რომ ეს მონაცემები ყოველდღიურად განიცდის მატებას, რაც პირდაპირ არის დამოკიდებული კორპუსზე მიერთებულ წყაროებში ახალი მასალების განთავსების ინტენსივობაზე. ყველა ეს საცდელი კორპუსი აღჭურვილია კლავიატურით (იხ., ქვემოთ, სურათი_8, სურათი_9, და სურათი_10), რაც მომხმარებელს საშუალებას აძლევს მისთვის სასურველი ნებისმიერი სიტყვით ან ფრაზით აწარმოოს ძიება შესაბამისი კორპუსის ბაზებში განთავსებულ ტექსტობრივ მონაცემებში [11].



სურათი 8 - ჩეჩნური ვებ კლავიატურა



სურათი 9 - ლეკური ვებ კლავიატურა



სურათი 10 - ყაბარდოული ვებ კლავიატურა

შემდგომისთვის, ჩვენ დაგეგმილი გვაქვს მსაგავსად აფხაზურისა ჩეჩნური, ყაბარდოული და ლეკური მკითხველი სისტემების შემუშავებაც (თუმცა, ჩვენ ამის გაკეთება ცხადია გაგვიჭირდება, ამ საკითხში ამ ენების სპეციალისტების ჩართულობის გარეშე). აგრეთვე, სამომავლოდ, ქართულთან ერთად ჩვენ (მხედველობაში გვაქვს კ.ფხაკაძის ხელმძღვანელობით მოქმედი ქართული ენის ტექნოლოგიების ცენტრის მკვლევართა ჯგუფი, რომლის წევრებიც არიან: დოქტორები - მ.ჩიქვინიძე და გ.ჩიჩუა, მაგისტრანტი - დ.კურცხალია და დოქტორანტები - კ.დემურჩევი, ნ.ოქროშიაშვილი, შ.მალიძე) დაგეგმილი გვაქვს აგრეთვე როგორც აფხაზური, ისე სხვა კავკასიური ენებისთვის ხმის ამომცნობი, ტექსტის მანაალიზებელი და ხმიდან-ხმაზე და ტექსტიდან-ტექსტზე მთარგმნელი სისტემების შემუშავების მიმართულებით უკვე მიმდინარე კვლევების შემდგომი გაღრმავება, რასაც ჩვენ ენობრივი ბარიერებისაგან თავისუფალი კავკასიის ფორმირებისა და, შესაბამისად, კავკასიური ენების ციფრული

კვდომის საფრთხისგან დაცვის მეტად მნიშვნელოვანი მიზნების ჭრილში განვიხილავთ [11].

2.4. ქართული ტიტრირებული მეტყველების კორპუსის საცდელი ვერსია

ქართული ტიტრირებული მეტყველების ამგები ინსტრუმენტების შემუშავების მიზნით შეიქმნა პროგრამული უზრუნველყოფა, სადაც შესაძლებელია მომხმარებლის რეგისტრაცია და რეგისტრაციის წარმატებით გავლის შემდეგ სისტემაში შესვლა. სარეგისტრაციო ფორმა <https://corpus.ge/Account/Registration> მისამართზე იხსნება. რეგისტრაციისას საჭიროა პიროვნებამ შეავსოს სარეგისტრაციო ფორმის აუცილებელი ველები. ესენია: სახელი, გვარი, საკონტაქტო იმეილი, დაბადების თარიღი, სქესი , მომხმარებლის სახელი და პაროლი (იხ., ქვემოთ, სურათი_11). რეგისტრაციის წარმატებით გავლის შემდგომ მომხმარებელი უნდა გადავიდეს <https://corpus.ge/Account/Login> მისამართზე განთავსებულ სისტემაში შესვლის გვერდზე და ახლა უკვე იქ უნდა მიუთითოს საკუთარი სარეგისტრაციო სახელი და პაროლი, რომელთა შემოწმებისა და ვალიდურობის დადგენის შემდგომ ის უკვე გადადის ე.წ. პირად კაბინეტში, სადაც მისთვის შესაძლებელი ხდება კორპუსის ბაზიდან შემთხვევითობის პრინციპით ამოიღოს იქ არსებული წინადადება, გაახმოვანოს ის მისთვის სასურველი ხარისხით და მერე უკვე შენახვის დილაკით ქართულ უნივერსალურ ჭკვიან კორპუსს მიაწოდოს მისი ხმით შექმნილი ქართული ტიტრირებული სამეტყველო მონაცემი, რის შემდეგაც ეს ახალი მონაცემი უერთდება ტიტრირებული სამეტყველო კორპუსის ფარგლებში უკვე დაგროვილ მონაცემებს (იხ., ქვემოთ, სურათი_12) [11].

Registration

UserName	<input type="text" value="Full Name"/>
FirstName	<input type="text" value="First Name"/>
LastName	<input type="text" value="Last Name"/>
Email	<input type="text" value="Email"/>
Date of Birth	<input type="text" value="mm/dd/yyyy"/>
Gender	<input type="radio"/> Female <input type="radio"/> Male
Password	<input type="text" value="Enter Your Password"/>
Confirm	<input type="text" value="Confirm Your Password"/>
<input type="button" value="Register"/>	

სურათი 11 - სარეგისტრაციო გვერდი

- ყოველთვის ახალი
- გახმოვანებულებიდან

მე იმ პერიოდში შეყვარებული მყავდა.

სურათი 12 - ქართული ტიტრირებული კორპუსი

გარდა ამისა, სისტემა მომხმარებელს საშუალებას აძლევს შემთხვევითობის პრინციპით აირჩიოს მის მიერვე ადრე უკვე გახმოვანებული წინადადება (ამისთვის მან უნდა მონიშნოს ალამი „გახმოვანებულებიდან“ (იხ., ზემოთ, სურათი 12)) ან აირჩიოს ახალი - მის მიერ ჯერ კიდევ გაუხმოვანებელი წინადადება (ამისთვის მან უნდა მონიშნოს ალამი „ყოველთვის ახალი“. ამასთან, ჩასაწერი წინადადების საბოლოო არჩევამდე შესაძლებელია შემოთავაზებული წინადადების გადარჩევა „ახალი წინადადება“ ღილაკზე დაჭერით (იხ., ზემოთ, სურათი_12)). ჩასაწერი წინადადების არჩევის შემდგომ ხორციელდება აუდიო მასალის ჩაწერა (პროცესი მოითხოვს კომპიუტერის აღჭურვას მიკროფონით). ჩაწერის დასაწყებად საჭიროა „ხმის ჩაწერა“ ღილაკზე მაუსის

დაჭერა, რის შემდგომაც გააქტიურდება მიკროფონი და დაიწყება ხმის ჩაწერა (იხ., ქვემოთ, სურათი_13 და სურათი_14) [11].



სურათი 13-14 - ტიტრირებული კორპუსში ხმის ჩაწერა

დროის მთვლელი ანიშნებს თუ რამდენი დროა გასული ხმის ჩაწერის დაწყებიდან. „შეწყვეტა“ ღილაკის დაჭერის შემთხვევაში ჩაწერა შეწყდება და ჩაწერილი ხმოვანი მასალა დროებით შეინახება. ჩაწერილი ხმოვანი მასალის დროებითი შენახვის შემდეგ მომხმარებელს (ანუ ამ შემთხვევაში ჩამწერს) საჭიროებიდან გამომდინარე მისი სურვილის ანუ გადაწყვეტილების შემთხვევაში შესაძლებლობა ეძლევა მოისმინოს მის მიერვე ახლად შექმნილი ჩანაწერი, გადაახვიოს ის, მოსმენისას დაარეგულიროს ხმა, გადმოიწეროს შექმნილი აუდიო ფაილი. ამასთან, თუ ჩანაწერის ხარისხის „გადასამოწმებელი“ ყველა ამ შესაძლებლობებით სარგებლობის შემდეგ მომხმარებელმა „უხარისხოდ“ მიიჩნია მის მიერვე წინასწარი სახით შემზადებული და დროებით შენახული ახალი ჩანაწერი, მას „წაშლა“ ღილაკის გააქტიურებით საშუალება ეძლევა წაშალოს ეს დროებით შენახული „უხარისხო მასალა“ და ხელმეორედ ჩაწეროს გასახმოვანებელი წინადადების ახალი - „უფრო ხარისხიანი“ ხმოვანი მეწყვილე. წინააღმდეგ შემთხვევაში, ანუ იმ შემთხვევაში თუ მომხმარებელმა დამაკმაყოფილებლად ჩათვალა დროებით შენახული ჩანაწერის „ხარისხი“, ის „შენახვა“ ღილაკის გააქტიურებით ამ დროებით შენახულ „ხარისხიან ჩანაწერს“ აგზავნის კორპუსის სერვერზე (იხ., ქვემოთ, სურათი_15) [11].



სურათი 15 - ხმის ჩამწერი ინსტრუმენტი

ამასთან, ეს სერვერზე უკვე გაგზავნილი მასალა შეინახება კორპუსის ტიტრირებული მეტყველების ბაზაში მისი ავტორის საიდენტიფიკაციო მონაცემებისა ჩაწერის დროის მინიშნებით. ამის შემდეგ, ეს ახალი უკვე სერვერზე შენახული ტიტრირებული ხმოვანი მონაცემი სერვერიდანვე მიუერთდება მომხმარებლის ხმოვანი ჩანაწერების ბაზას, რომელთანაც მომხმარებელს უშუალო წვდომა აქვს, რაც მას სურვილის შემთხვევაში თავისი ჩანაწერის მოსმენისა და ჩამოტვირთვის საშუალებებს აძლევს (იხ., ქვემოთ, სურათი_16) [11].

ჩანაწერები		
#	ტექსტი	თარიღი
1	ეს პერიოდი გაგრძელდა 3 წელი	09-12-2017 17:24:28
2	სამსახურში როგორ სხვა დაწარჩუნებს ისე მოვეცევი.	09-12-2017 14:26:17
3	სამშრომლები მიმდინარეობს და 17 აგვისტოს დასრულდება	09-12-2017 14:03:26
4	-თავი ვერ შევიკავე, ძალიან შენეშინდა.	09-12-2017 13:57:48

სურათი 16 - ჩაწერილი ხმები

გარდა ამისა, ტიტრირებული მონაცემების სისტემა მომხმარებელს საშუალებას აძლევს ერთიდაიგივე წინადადებაზე მიბმით შეინახოს მისი სხვადასხვა გახმოვანებული ვერსია. ამასთან, შენახვისა და მომხმარებლის დაბადების თარიღების მეშვეობით, სისტემა ჩადგმული ლოგიკით მარტივად ითვლის თუ რა ასაკის იყო მომხმარებელი, როცა მან ესა თუ ის კონკრეტული ჩანაწერი გააკეთა. სისტემით უზრუნველყოფილი ეს შესაძლებლობა მეტყველების ამოცნობისა და სინთეზის ამოცანებში შედეგების წინ წაწევის გარდა, მნიშვნელოვანია აგრეთვე სამეტყველო ხმის ასაკობრივ ცვლილებებზე სამეცნიერო დაკვირვებების საწარმოებლად.

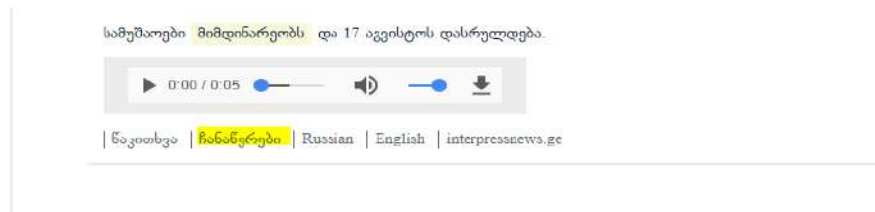
დღეს-დღეობით მომხმარებლის გახმოვანებული მასალის ნახვა შესაძლებელია ჩვენი კორპუსის მთავარი გვერდიდან ქართულ ენაზე გადასვლითა და საძიებო ველში წინადადების ან სიტყვის ჩაწერისას “Rec” ალამის მონიშნით წამოწყებული ძეხნით (იხ., ქვემოთ, სურათი_17) [11].



სურათი 17 - ძებნა ტიტრირებულ მონაცემებში

მომეზნილ მასალას დაბლა ზოლში თან ერთვის ღილაკი „ჩანაწერები“, მასზე დაჭერის შემთხვევაში ხილვადი ხდება ამ წინადადების გახმოვანებული ყველა ვერსია და შესაძლებელია მათი მოსმენა (იხ., ქვემოთ, სურათი 18).

ფრაზა: "მიმდინარეობს" | რაოდენობა: 124738 | ფარდობითი სიხშირე: 0.0447814122%



სურათი 18 - ტიტრირებული მონაცემების გამოტანა

2.5. აფხაზური, ჩეჩნური, ლეკური და ყაზარდოული ტიტრირებული მეტყველების კორპუსების ამგები ინსტრუმენტები

ტიტრირებული მეტყველების კორპუსის ამგები ზემოთ წარმოდგენილი სისტემა ჯერ-ჯერობით მორგებულია მხოლოდ ქართულ კორპუსზე და, შესაბამისად, ამ ინსტრუმენტით ჯერ მხოლოდ ქართული ტიტრირებული მეტყველების კორპუსი იგება. თუმცა, ამ სისტემის სხვადასხვა ენობრივი ინდექსაციით ჩვენ უკვე საცდელი სახით შემუშავებული გვაქვს ამ სისტემის სხვა ენობრივი ვერსიებიც, რაც გასაგებს ხდის, რომ ჩვენ უკვე მზადავართ აფხაზური, ჩეჩნური, ლეკური, და ყაზარდოული, ენების ტიტრირებული სამეტყველო კორპუსების ასაგებად, რაც, ცხადია, მეტად მნიშვნელოვანი ამოსავალი მოცემულობა იქნება ამ ენების ციფრული კვდომის საფრთხისგან დაცვის თვალსაზრისით (იხ.,

ქვემოთ, სურათი_19, სურათი_20, სურათი_21, სურათი_22, სურათი_23, სურათი_24, სურათი_25) [11].



ჩანაწერები

#	თაჩსტი	თარიღი	დრო	აუდიო
1	ეს პერიოდი გაგრძელდა 3 წელი.	09-12-2017 17:24:28	▶ 0:00 / 0:03	🔊 🔇 📄
2	სამსახურში როდრე სხვა დანარჩენებს ისე მოგვამცდვი.	09-12-2017 14:26:17	▶ 0:00 / 0:03	🔊 🔇 📄

სურათი 19 - ტიტრიებულ კორპუსში ენების მხარდაჭერა

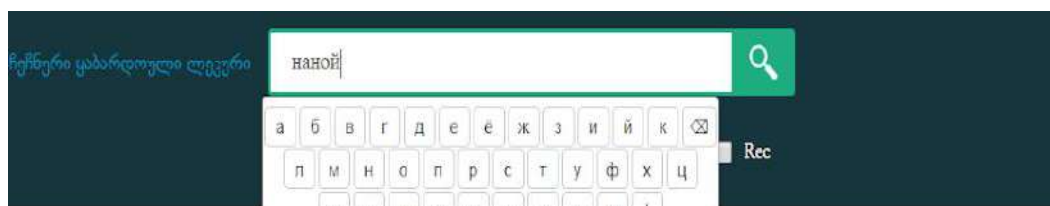


ფრაზა: "наной" | რაოდენობა: 1 | ფარდობითი სიხშირე: 3.555e-7% | შემოსვლის დრო: 2018-02-11 09:16:13

Духанберш, наной .су. реза бецури

ნაკითხვა | ჩანაწერები | Russian | English | radiomarsho.com

სურათი 20 - ჩეჩნური სიტყვის ძებნა



ფრაზა: "наной" | რაოდენობა: 1 | ფარდობითი სიხშირე: 3.555e-7% | შემოსვლის დრო: 2018-02-11 09:16:13

Духанберш, наной .су. реза бецури



ნაკითხვა | ჩანაწერები | Russian | English | radiomarsho.com

სურათი 21 - ტიტრიებული ჩანაწერის ჩვენება სიტყვაზე наной



ფრაზა: "азмет" | რაოდენობა: 3 | ფარდობითი სიხშირე: 0.0000010664% | შემოსულის დრო: 2561-02-19 10:35:46

,- отметил Азмет Басте.

| ნაკობეა | ჩანაწერები | Russian | English | cir.rus4all.ru

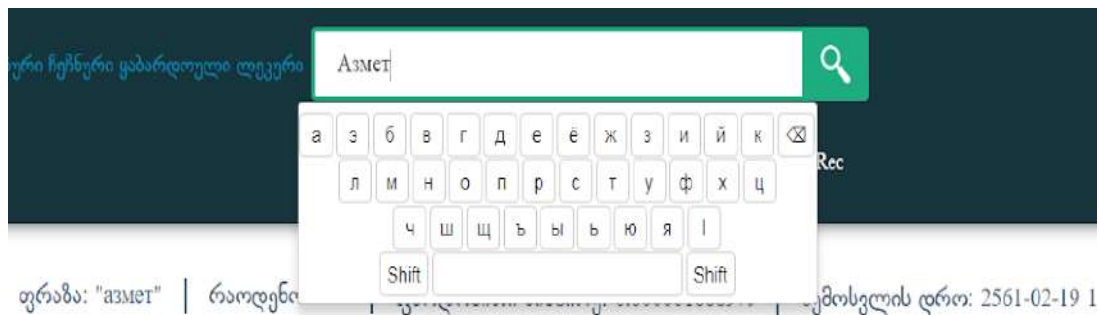
Как отметил директор ансамбля Азмет Басте, гастрольный график практически утвержден, в него вошло около 85 концертов по стране и за рубежом.

| ნაკობეა | Russian | English | cir.rus4all.ru

, и Азмет Малиш - заслуженный тренер СССР, мастер спорта по дзюдо.

| ნაკობეა | Russian | English | cir.rus4all.ru

სურათი 22 - ყაზარდოული სიტყვის ძებნა



ფრაზა: "азмет" | რაოდენობა: 3 | ფარდობითი სიხშირე: 0.0000010664% | შემოსულის დრო: 2561-02-19 10:35:46

,- отметил Азмет Басте.

▶ 0:00 / 0:03 🔊 ⬇️

| ნაკობეა | ჩანაწერები | Russian | English | cir.rus4all.ru

Как отметил директор ансамбля Азмет Басте, гастрольный график практически утвержден, в него вошло около 85 концертов по стране и за рубежом.

| ნაკობეა | Russian | English | cir.rus4all.ru

, и Азмет Малиш - заслуженный тренер СССР, мастер спорта по дзюдо.

| ნაკობეა | Russian | English | cir.rus4all.ru

სურათი 23 - ტიტრირებული ჩანაწერის ჩვენება სიტყვაზე Азмет


 კავასიური ენები: ავსტრული ჩქნური დაბარდლოლი ლეკური

სულ შემოსული: 401389
 განსხვავებული: 130483
 ბოლო დღეს: 0

სიტყვების რაოდენობა: -დან -მდე Rec

ფრაზა: "кассалий" | რაოდენობა: 1 | ფარდობითი სიხშირე: 3.555e-7% | შემოსვლის დრო: 2018-02-19 13:03:54

Кассалий бахаййеса иа билет бахъаг.

[წაკითხვა](#) | [ჩანაწერები](#) | [Russian](#) | [English](#) | [ilchi.info](#)

სურათი 24 - ლეკური სიტყვის ძებნა

რი ჩქნური დაბარდლოლი ლეკური

ფრაზა: "кассалий" | რაოდენობა: 1 | ფარდობითი სიხშირე: 3.555e-7% | შემოსვლის დრო: 2018-02-19 13:03:54

ავსტრული ჩქნური დაბარდლოლი ლეკური

აბ ვ გ დ ე ё ж з и й к
 л м н о п р с т у ф х ц
 ч ш щ ъ ы ь э ю я

Shift Shift

Кассалий бахаййеса иа билет бахъаг.

0:00 / 0:06

[წაკითხვა](#) | [ჩანაწერები](#) | [Russian](#) | [English](#) | [ilchi.info](#)

სურათი 25 - ტიტრირებული ჩანაწერის ჩვენება სიტყვაზე Кассалий

რი ჩქნური დაბარდლოლი ლეკური

სიტყვების რაოდენობა: -დან -მდე Rec

ფრაზა: "вари" | რაოდენობა: 70 | ფარდობითი სიხშირე: 0.0000248785% | შემოსვლის დრო: 2018-02-18 13:03:54

Вари ттул даврит мурад пахру.

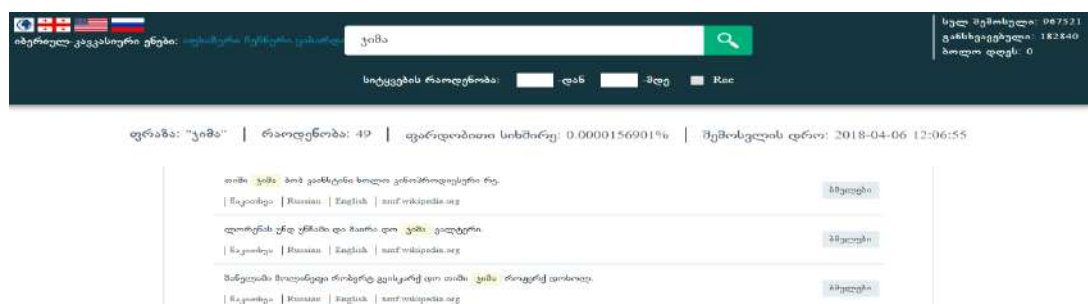
0:00 / 0:05

[წაკითხვა](#) | [ჩანაწერები](#) | [Russian](#) | [English](#) | [ilchi.info](#)

სურათი 26 - ტიტრირებული ავსტრული ჩანაწერი სიტყვაზე Вари

2.6. ქართულ უნივერსალურ ჭკვიან კორპუსში ჩადგმული მეგრული თვითგანვითარებადი კორპუსი

ქართული ენის ტექნოლოგიების ცენტრის თვითგანვითარებადი ქართული ჭკვიანი ვებ-კორპუსის საცდელ-სამომხმარებლო ვერსიას აფხაზური, ჩეჩნური, ყაზახური და ლეკური ენების კორპუსებთან ერთად სადოქტორო კვლევის ფარგლებში საცდელი სახით დაემატა აგრეთვე მეგრული ენის კორპუსიც. ამგვარად, ახლა უკვე შესაძლებელია ჩვენს უნივერსალურ კორპუსში მოიძებნოს მასალა მეგრულ ენაზეც. ძიება შესაძლებელია როგორც სიტყვის, ასევე ფრაზის მეშვეობითაც. ქართულ ჭკვიან თვითგანვითარებად ვებ-კორპუსში ცალკე კორპუსად ჩადგმული მეგრული ენის კორპუსის განსავითარებლად ცხადია საჭიროა მეგრულენოვანი წყაროები, რაც პერმანენტულად და, ასევე, ავტომატურად ამ წყაროებიდან მეგრულენოვანი ტექსტების კორპუსში შემოტანის საშუალებას იძლევა.¹⁶ დღეისათვის მოძიებულია და კორპუსზე ავტომატურ რეჟიმში უკვე მიერთებული რიგი მეგრულენოვანი წყაროები. ამასთან, ამ წყაროებიდან სულ ჯამურად უკვე წამოღებულია 967521 მეგრული სიტყვა, რომელთაგან 182840 განსხვავებულია (მონაცემები აღებულია 2019 წლის 18 დეკემბრის ჩვენების მიხედვით). როგორც აღინიშნა ეს მონაცემები ყოველდღიურად განიცდის მატებას, რაც დამოკიდებულია წყაროებში ახალი მასალის განთავსების ინტენსივობაზე.



სურათი 27 - ძებნა მეგრულ კორპუსში

¹⁶ კორპუსში მეგრული ტექსტები შემოდის ისე, როგორც ეს უკვე აღწერილია აფხაზური და სხვა კავკასიური ენებისთვის. ამიტომ აქ ამ პროცედურას აღარ აღვწერთ.

სურათზე ნაჩვენებია „ჯიმა“ სიტყვით წამოწყებული ძეგლის შედეგი. მეგრულისთვის რეალიზებულია დამებნილი ფრაზის ავტომატური კითხვა.

ფრაზა: "ჯიმა" | რაოდენობა: 49 | ფარდობითი სიხშირე: 0.0000156901% | შემოსვლის დრო: 2018-04-06 12:06:55

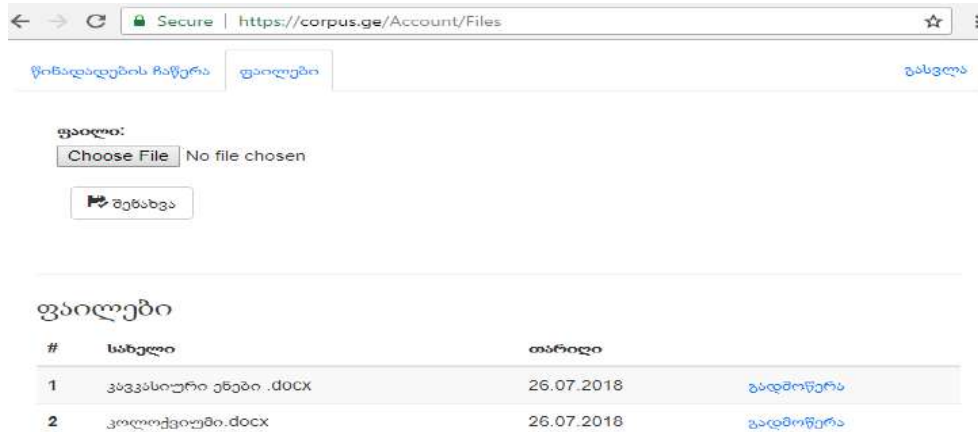
სურათი 28 - ინფორმაცია სიტყვაზე „ჯიმა“

სურათზე გამოტანილია „ჯიმა“ სიტყვის სიხშირე ანუ რაოდენობა, ფარდობითი სიხშირე და კორპუსში ამ სიტყვის შემოსვლის თარიღი. ამგვარად, შეიძლება ითქვას, რომ საცდელი სახით უკვე აგებულია მეგრული თვითგანვითარებადი ჭკვიანი კორპუსი, რომელშიც საცდელი სახით უკვე ფუნქციონირებს მეგრული მკითხველი, რაც, მთლიანობაში, ცხადია, არაუმნიშვნელო საწყისი მოცემულობაა მეგრული ენის ციფრული კვდომის საფრთხისგან დაცვისა და ტექნოლოგიური განვითარების თვალსაზრისით.

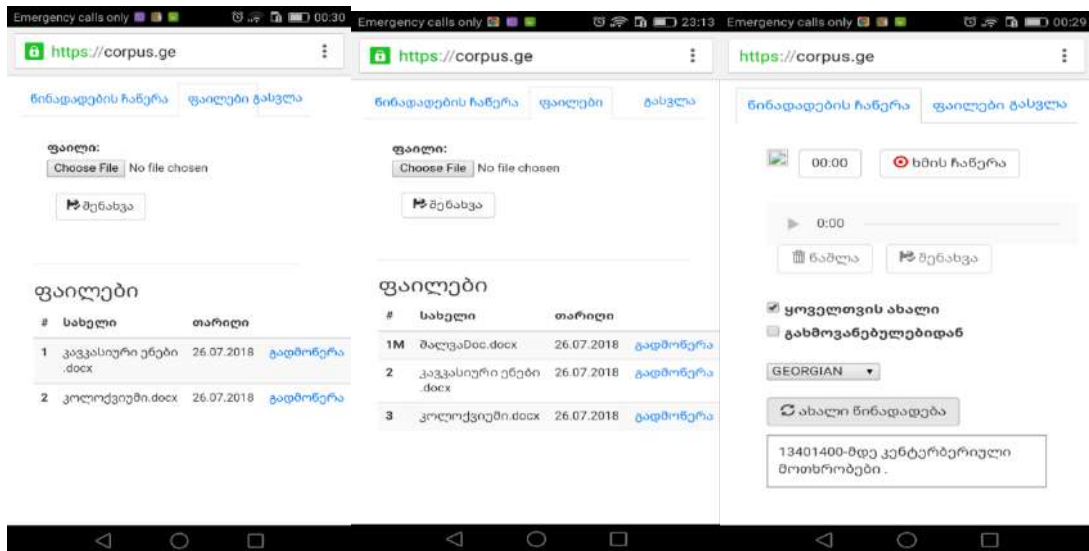
2.7. მომხმარებლების მოწყობილობებიდან (კომპიუტერი, მობილური) კორპუსის გამაძლიერებელი ინსტრუმენტები

სადოქტორო კვლევის ფარგლებში შემუშავდა ჩვენი კორპუსის მომხმარებლის მოწყობილობებიდან გამაძლიერებელი ინსტრუმენტები. კერძოდ, სისტემა რეგისტრირებულ მომხმარებელს აძლევს საშუალებას: 1. თავისი კომპიუტერიდან/მობილურიდან მიუერთოს კორპუსს ვორდის ელექტრონული დოკუმენტი; 2. ჩამოტვირთოს კორპუსზე მის მიერვე მანამდე უკვე მიერთებული ვორდის დოკუმენტი. ვორდის ფაილების ამტვირთავის გვერდზე მომხმარებელს შეუძლია მონიშნოს თავისთვის სასურველი ფაილი კომპიუტერში და აიტანოს ჩვენ სისტემაში, სადაც ვორდის ფაილიდან ამოიკითხება ტექსტი და ჩვენ სისტემაში გადავა ჩაშლილი სახით. ანუ იმ სახით როგორ ეს ზემოთ არის აღწერილი ქართული ენის კორპუსისათვის, რაც, ცხადია, კორპუსს, როგორც მოწესრიგებულ მონაცემთა ბაზას, აძლიერებს. ქვემოთ, სურათ_29-ზე ნაჩვენებია როგორ

ატვირთა მომხმარებელმა კომპიუტერიდან ვორდის დოკუმენტის ტიპის ორი ფაილი. ამასთან, როგორც ზემოთაც აღინიშნა, ვორდის ფაილების ამტვირთავს აქვს მობილური ვერსიაც ანუ მობილურის მხარდაჭერაც (იხილეთ, ქვემოთ, სურათი_30 და სურათი_31).



სურათი 29 - გამაძლიერებელი ინსტრუმენტი კომპიუტერისთვის

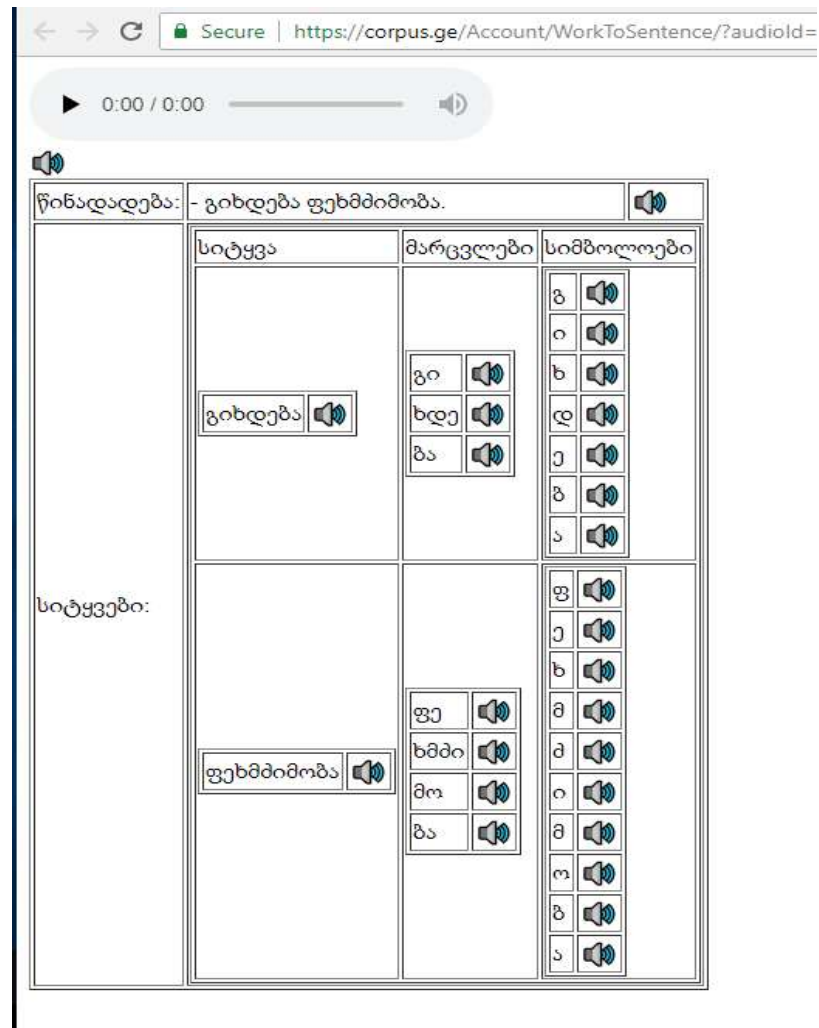


სურათი 30-31-32 - გამაძლიერებელი ინსტრუმენტი მობილურისთვის

სურათ_32-ზე მითითებულია მობილურიდან ატვირთული ფაილი „შალვაDoc.docx“ რომელსაც რეგითობაში აქვს აღნიშვნა როგორც მობილურიდან ატვირთული („M“). ასევე, სისტემას ჩაემატა შესაძლებლობა რომ მობილური ტელეფონებიდანაც შეძლოს მომხმარებელმა ხმის ჩაწერა ტიტრირებული სამეტყველო მონაცემების კორპუსში. - იხილეთ, ზემოთ, სურათი_32.

2.8. ქართული და აფხაზური ტიტრირებული სამეტყველო მონაცემების სეგმენტატორების საცდელი სისტემები

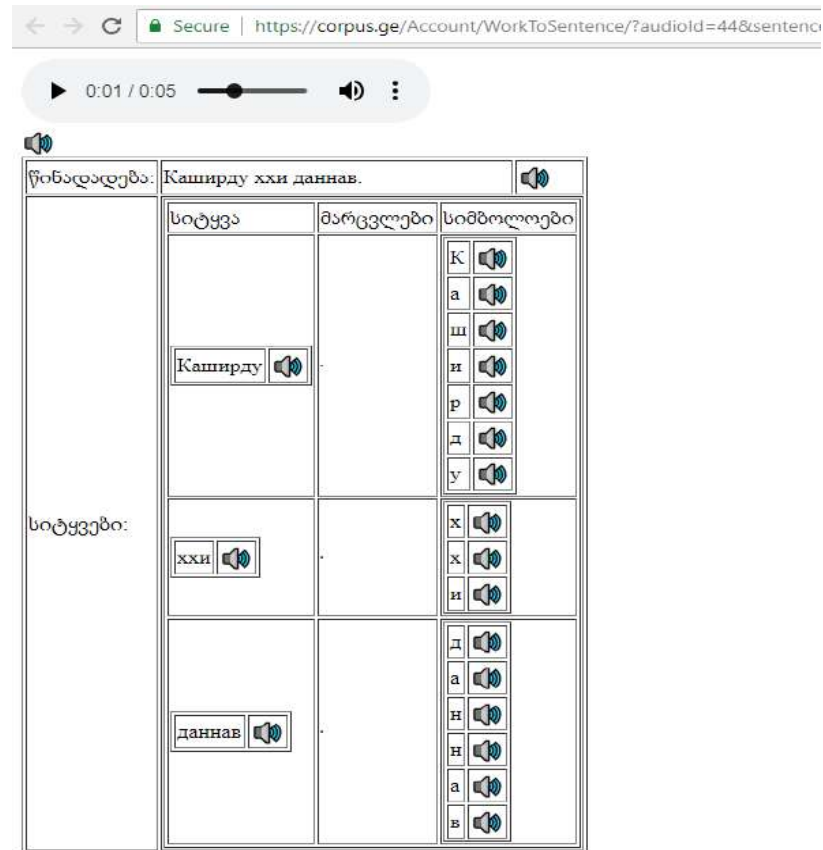
იხილეთ, ქვემოთ, სურათი 33.



სურათი 33 - ქართული სეგმენტატორი კომპიუტერისთვის

როგორც სურათიდან ჩანს, სეგმენტატორმა დასამუშავებელი ტიტრირებული სამეტყველო მონაცემი ჩაშალა სიტყვებად, სიტყვები - მარცვლებად, მარცვლები - ასოებად. ამასთან, სისტემა ამას აკეთებს ისე, რომ თითოეულ ფრაგმენტს უწყვილებს სამეტყველო მონაცემის შესაბამის ფრაგმენტს. ანუ: 1. წინადადების ჩაშლის შედეგად წარმოქმნილ სიტყვას უწყვილდება ამ სიტყვის სამეტყველო შესატყვისი; 2. სიტყვის ჩაშლის შედეგად წარმოქმნილ მარცვალს/ასოს უწყვილდება ამ მარცვლის/ასოს

სამეტყველო შესატყვისი. ანალოგიურად მუშაობს სეგმენტატორი აფხაზურ ენაზეც უბრალოდ ამ შემთხვევაში დამარცვლა ვერ ხერხდება (იხილეთ, ქვემოთ, სურათი 34).



სურათი 34 - აფხაზური სეგმენტატორი კომპიუტერისთვის

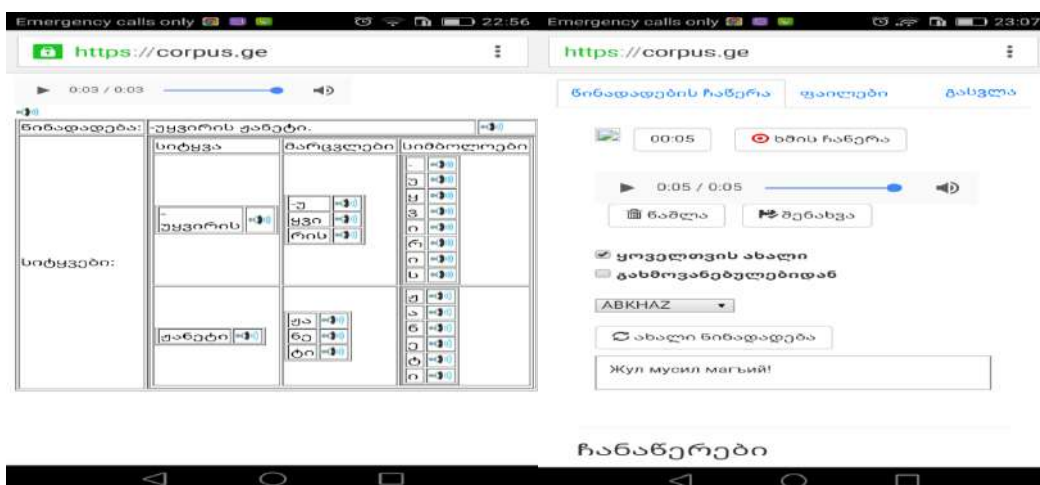
კვლევის ამ ეტაპზე სისტემა სრული სიზუსტით ვერ მუშაობს, რაც განპირობებულია იმით, რომ სეგმენტირებისას ჩვენ ჯერ კიდევ არ ვითვალისწინებთ შემადგენელი ასო-ბგერების ბუნებრივ ფიზიკურ თავისებურებებს. თუმცა, მიუხედავად ამისა, თანაბრად წარმოთქმულ მცირე ზომის ფრაზებსა და წინადადებებში სისტემა საკმარისი სიზუსტით აკეთებს სეგმენტირებას ანუ თავდაპირველი ტიტრირებული მოცემულობის ზემოაღწერილი სახის საკმაოდ კომპლექსურ დაჭრას სხვადასხვა დონის შემადგენელ ერთეულოვან ნაწილებად. ხმის ჩაწერა სისტემაში

შესაძლებელია ასევე მობილურიდანაც. ქვემოთ მოცემულია მაგალითი სადაც ხდება ქართული წინადადების ჩაწერა მობილურის დახმარებით (იხილეთ, ქვემოთ, სურათი_35).



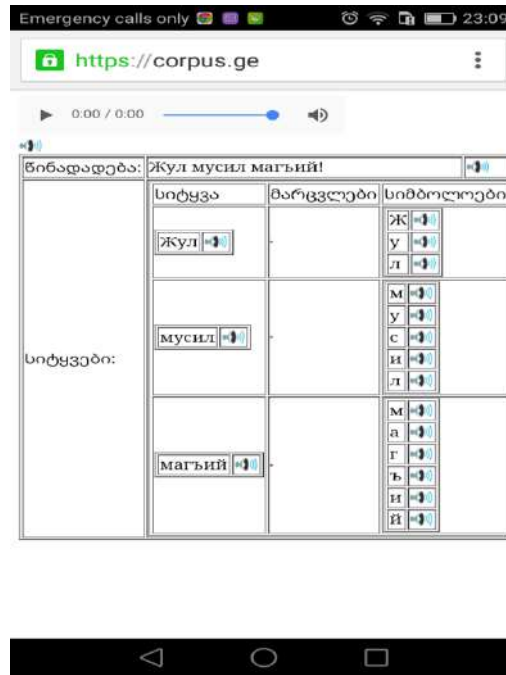
სურათი 35 - ქართული სეგმენტატორი მობილურისთვის

როგორ ეს ზემოთ მოცემული სურათიდანაც ჩანს, სისტემამ ხმა ჩაიწერა და შეინახა. ახლა ვნახოთ როგორ ჩაშლის სეგმენტატორი მობილურიდან ამ უკვე ჩაწერილ ხმას (იხილეთ, ქვემოთ, სურათი 36).



სურათი 36-37 - აფხაზური სეგმენტატორი მობილურისთვის

სურათ_38-ზე დოკუმენტირებულია აფხაზური წინადადების ჩაწერა მობილურიდან. ქვემოთ, სურათ_38-ზე კი წარმოდგენილია მობილურიდან ჩაწერილი აფხაზური ტიტრირებული სამეტყველო მონაცემის სეგმენტაციის მაგალითი.



სურათი 38 - აფხაზური სეგმენტაციის მაგალითი

2.9. ქართული ჭკვიანი ფურცლის პირველი საცდელი ორენოვანი (ქართულ-აფხაზური) ვერსია (ორენოვანი ბეჭედა; ორენოვანი კითხვა; ორენოვანი ხმოვანი მართვა; ორენოვანი ორთოგრაფიული მართლწერა; ქართული სინტაქსური მართლწერა)

სადისერტაციო კვლევის ფარგლებში აიგო კორპუსში ჩადგმული ქართული თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსიის თითქმის სრული სამომხმარებლო ვერსია და ქართული თვითგანვითარებადი გრამატიკული მართლწერის საცდელი ვერსიის საცდელ-სამომხმარებლო ვერსია. გარდა ამისა, სადისერტაციო კვლევის ფარგლებში აიგო აფხაზური თვითგანვითარებადი ორთოგრაფიული მართლწერის პირველი საცდელ-სამომხმარებლო ვერსია. ასევე, როგორც ზემოთ აღინიშნა, აიგო აგრეთვე კორპუსში ჩადგმული

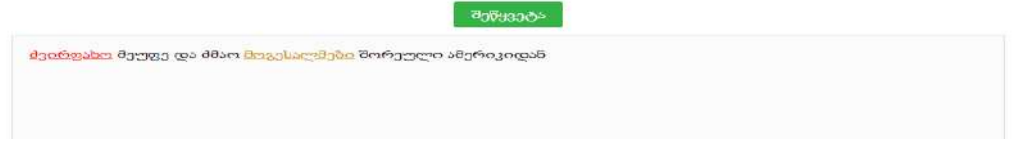
აფხაზური მკითხველის პირველი და ჯერ-ჯერობით ერთადერთი საცდელი ვერსია. ამან, მთლიანობაში, საშუალება მოგვცა კორპუსში ჩადგმული სახით არსებული ქართული ჭკვიანი ფურცელი გაგვეფართოვებინა აფხაზური ენით და ამ გზით შეგვექმნა ქართული ჭკვიანი ფურცლის ორენოვანი (ქართულ-აფხაზური) საცდელი ვერსია, რომელშიც ფუნქციონირებს ორენოვანი ანუ ქართულ-აფხაზური ბეჭედა; ორენოვანი ანუ ქართულ-აფხაზური კითხვა; ორენოვანი ანუ ქართულ-აფხაზური ხმოვანი მართვა; ორენოვანი ანუ ქართულ-აფხაზური ორთოგრაფიული მართლწერა და, აგრეთვე, ქართული სინტაქსური მართლწერა. ამასთან, თავიდანვე აღვნიშნავთ, რომ ფხაკაძე-ჩიქვინიძის მეთოდებზე დაყრდნობით ჩვენს მიერ საცდელი სახით უკვე აგებული აფხაზური ორთოგრაფიული და ქართული სინტაქსური მართლმწერი, ისევე როგორც ფხაკაძე-ჩიჩუას მეთოდებზე დაყრდნობით აგებული აფხაზური მკითხველი და აფხაზური მეტყველების ამომცნობი, ჩვენს ხელთ არსებული მონაცემებით, პირველი და ჯერ-ჯერობით ერთადერთი ამ ტიპის ქართულ - აფხაზური ენობრივი ტექნოლოგიებია. - ქვემოთ უფრო დეტალურად უკვე აღნიშნულზე.

2.9.1. კორპუსში ჩადგმული ქართული თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსიის თითქმის სრული სამომხმარებლო ვერსია

აქ, სურათ_39-ზე, შუა ველი განკუთვნილია ტექსტის შესაყვანად, რის შემდეგაც შემოწმება დილაკზე დაჭერით სისტემა ანალიზებს რომელი სიტყვაა არასწორი ანუ მცდარი ფორმის და აფერადებს მას წითლად.



სურათი 39 - ქართული ორთოგრაფიული მართლმწერი (1)



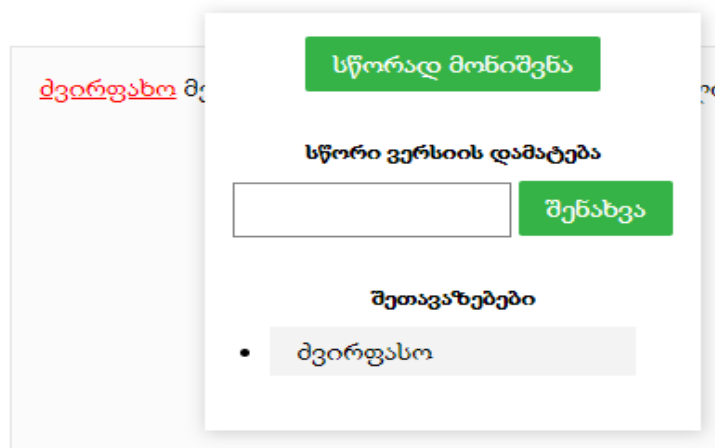
სურათი 40 - ქართული ორთოგრაფიული მართლმწერი (2)

ასევე შესაძლებელია ფურცელზე არსებული ქართული წინადადების მოსმენა „წაკითხვა“ ლილაკზე დაჭერით ან/და შესაბამისი ხმოვანი ბრძანების მიცემით იწყება დაკვეთილი კითხვის პროცესი, რის შემდეგაც ლილაკი ავტომატურად გარდაიქმნება „შეყვება“ ლილაკად (იხილეთ, ზემოთ, სურათი_40), რომლის გააქტიურებითაც ან/და შესაბამისი ხმოვანი ბრძანების მიცემით უკვე კითხვა შეგვიძლია შევწყვიტოთ ნებისმიერ ჩვენთვის სასურველ დროს, რის შემდეგაც „შეწყვეტა“ ლილაკი კვლავ გარდაიქმნება „წაკითხვა“ ლილაკად (იხილეთ, ქვემოთ, სურათი_41).



სურათი 41 - ქართული ორთოგრაფიული მართლმწერი (3)

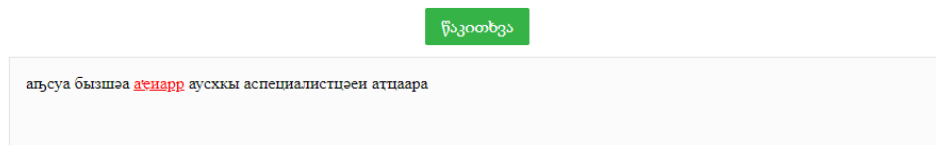
გარდა ამისა, გაწითლებულ სიტყვაზე დაჭერით სისტემა გვამღევს საშუალებას არასწორი სიტყვა ჩავასწოროთ ბაზაში არსებული სწორი სიტყვით თუ რა თქმა უნდა ასეთი მიახლოებითი არჩევანი გათვალადა (იხილეთ, ქვემოთ, სურათი_42).



სურათი 42 - ქართული ორთოგრაფიული მართლმწერის კორექტორი

2.9.2. კორპუსში ჩადგმული აფხაზური თვითგანვითარებადი ორთოგრაფიული მართლმწერის საცდელ-სამომხმარებლო ვერსია

ხაზგასამელია, რომ უკვე აგებული ქართული ჭკვიანი ფურცლის ორენოვანი ვერსია, რომლის ქართულენოვანი ფუნქციონირება უკვე მოკლედ აღვწერეთ, ანალოგიურადვე ფუნქციონირებს აფხაზურ ენაშიც.



სურათი 43 - აფხაზური ორთოგრაფიული მართლმწერი (1)

ანუ, სისტემაში ჩაშენებულია აფხაზური კითხვაც (იხილეთ, ზემოთ, სურათი_43) და, ასევე, აფხაზური კლავიატურა (იხილეთ, ქვემოთ, სურათი_44).



სურათი 44 - აფხაზური ორთოგრაფიული მართლმწერი (2)

ასევე, ჩაშენებულ აფხაზურ მკითხველში, მსგავსად ქართული მკითხველისა, „წაკითხვა“ ღილაკზე დაჭერით და/ან შესაბამისი ხმოვანი ბრძანების მიცემით იწყება კითხვის პროცესი და ღილაკი გარდაიქმნება „შეწყვეტა“ ღილაკად და პირიქით (იხ., სურათი_45 და სურათი_46).

შეწყვეტა

ауьсуа бызшәа **аҕһаар** ауьхкы аспециалисттәеи аттаара

სურათი 45 - ავხაზური ორთოგრაფიული მართლმწერი (3)

წაკითხვა

ауьсуа бызшәа **аҕһаар** ауьхкы аспециалисттәеи аттаара

სურათი 46 - ავხაზური ორთოგრაფიული მართლმწერი (4)

2.9.3. კორპუსში ჩადგმული ქართული თვითგანვითარებადი გრამატიკული მართლმწერის საცდელი ვერსიის საცდელ-სამომხმარებლო ვერსია

როგორც ზემოთაც აღინიშნა ქართული ჭკვიანი ფურცლის ორენოვანი (ქართულ-ავხაზური) საცდელი ვერსიაში ზემოთ მოკლედ უკვე აღწერილი ორენოვანი ბეჭდვის, კითხვის, ხმოვანი მართვისა და ორთოგრაფიული მართლმწერის გარდა საცდელ-სამომხმარებლო სახით რეალიზებულია აგრეთვე თვითგანვითარებადი ქართული სინტაქსური მართლმწერის ანუ სინტაქსური ანალიზის სისტემა. თვითგანვითარებადი სინტაქსური ანალიზის ამ სისტემაში გვხვდება ტექსტური ველი სადაც სასურველი წინადადება უნდა შევიყვანოთ და ასევე ღილაკი „ანალიზი“ რაზე დაჭერისთანავე სისტემა გააკეთებს ტექსტურ ველში შეყვანილი წინადადების სინტაქსური ანალიზს. სისტემის საწყისი გვერდი გამოიყურება შემდეგნაირად (იხ. სურათი_47):

სინტაქსური ანალიზი

ანალიზი

სურათი 47 - სინტაქსური ანალიზის საწყისი გვერდი

ამასთან, თუ ტექსტურ ველში ჩავწერთ წინადადებას სისტემა კორპუსზე დაყრდნობით თვითონვე გაანალიზებს წინადადებას, ანუ თითოეულ სიტყვას თავადვე მიცემს თავის ტიპს, თავადვე მოძებნის ზმნას და თავადვე ამოიღებს ზმნის სინტაქსურ ფორმულას, რომელსაც დაგვიწერს სრულად და წინადადების ამ ფორმულაზე ანუ, რაც იგივეა, წინადადების ამ ზოგად სინტაქსურ ფორმაზე დაყრდნობით დაიწყებს შემოსული წინადადების სინტაქსურ გაანალიზებას. სურათ_48-ზე მოცემულია მაგალითი სადაც გასაანალიზებელი წინადადებაა „ბიჭი წერს“.

სინტაქსური ანალიზი

ანალიზი

წინადადება სწორია

სიტყვები:

1. ბიჭი - არსებითი სახელი Ni ბრუნვაში
2. წერს - ზმნა

სინტაქსური ფორმულა:

Ni [V (ის/Nი) [მას/Nს|მათ/Nებს] [Nს/Nებს]]

1. [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]
2. ბიჭი [მას/Nს|მათ/Nებს] [Nს/Nებს]

სურათი 48 - სინტაქსური ანალიზი მოქმედებაში (1)

როგორც სურათიდან ჩანს, სისტემამ გაანალიზა თითოეული სიტყვა და გვითხრა რომ „ბიჭი“ სინტაქსურად არის Ni ტიპის არსებითი სახელი. ასევე, „წერს“ სიტყვაზე გვითხრა რომ სინტაქსურად არის „[V (ის/Nი) (მას/Nს|მათ/Nებს) (Nს/Nებს)]“ ტიპის ზმნა, რაც იმას ნიშნავს, რომ ეს ზმნა არის სამ ადგილიანი ფორმა, რომლის პირველი ადგილი შეიძლება დაკომპლექტდეს „ის“ ან Ni ტიპის სახელით, მეორე ადგილი - „მას“ ან Nს

ტიპის სახელით ანდა „მათ“ ან Nებს ტიპის სახელით, მესამე ადგილი - Nს ან Nებს ტიპის სახელით. ასევე, სისტემამ მოგვცა გასაანალიზებელი წინადადების სინტაქსური ტიპი ანუ ფორმულა, რომელიც ჩაიწერება ანუ მოიცემა $N_i [V (ის/N_i) (მას/Nს|მათ/Nებს) (Nს/Nებს)]$ გამოსახულებით. ამის შემდეგ სისტემამ ზმნის სინტაქსურ ფორმულაში წინადადების შემადგენელი სიტყვა „ბიჭი“, რომელიც მას უკვე ამოცნობილი აქვს როგორც N_i ტიპის არსებითი სახელი, ჩასვა პირველ ანუ (ის/N_i) ტიპის ადგილზე და შედეგად მოგვცა „[ბიჭი V (მას/Nს|მათ/Nებს) (Nს/Nებს)]“ ფორმულა. ამგვარად, რადგან წინადადებაში სხვა სიტყვა არ არის, სისტემამ ანალიზის პროცესი დაასრულა და დაასკვნა რომ წინადადება სწორია.

ახლა ვნახოთ სისტემა რას გვეტყვის წინადადების „ბიჭმა წერს“ ანალიზის შემთხვევაში (იხ. სურათი_49).

სინტაქსური ანალიზი

ანალიზი

წინადადება არასწორია

სიტყვა 'ბიჭმა' ბრუნვა არ შეესაბამება ზმნას შესაძლო ფორმები: ბიჭი, ბიჭს

სიტყვები:

1. ბიჭმა - არსებითი სახელი Nმა ბრუნვაში
2. წერს - ზმნა

სინტაქსური ფორმულა:

Nმა [V [ის/N_i] [მას/Nს|მათ/Nებს] [Nს/Nებს]]

1. [ის/N_i] [მას/Nს|მათ/Nებს] [Nს/Nებს]

სურათი 49 - სინტაქსური ანალიზი მოქმედებაში (2)

სისტემამ წინადადება შეგვიფასა როგორც არასწორი - დაასკვნა რომ არსებითი სახელი ბრუნვაში არ ეთანხმება ზმნას. ასევე შემოგვთავაზა შემჩნეული სინტაქსური შეცდომის გამასწორებელი ვერსიები. კერძოდ, გვითხრა რომ „ბიჭმა“ არსებითი სახელი, რომელიც Nმა ფორმაშია, არ ეთანხმება ზმნას და გვითხრა ისიც, რომ ამ სიტყვის შესაძლო სწორი ფორმები შეიძლება იყოს „ბიჭი“ ან „ბიჭს“ .

ასევე აანალიზებს სისტემა ზმნასთან სახელის რიცხვში არათავსებადობას. მაგალითად, გაანალიზა რა სისტემამ წინადადება „ბიჭები წერს“, დაასკვნა რომ „ბიჭები“ რიცხვში არ ეთანხმება ზმნას და გვითხრა, რომ შესაძლო სწორი ფორმაა „ბიჭი“ (იხ. სურათი_50).

სინტაქსური ანალიზი

ბიჭები წერს ანალიზი

წინადადება არასწორია

სიტყვა 'ბიჭები' რიცხვში არ ეთანხმება, უნდა იყოს მხოლოდით ფორმაში 'ბიჭი'

სიტყვები:

- ბიჭები - არსებითი სახელი Nები ბრუნვაში
- წერს - ზმნა

სინტაქსური ფორმულა:

Nები [V [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]]

- [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]

სურათი 50 - სინტაქსური ანალიზი მოქმედებაში (3)

სისტემა ასევე ასწორებს ანუ აანალიზებს სახელის ზმნასთან რიცხვში და ბრუნვაში ერთდროული არათანხვედრის შემთხვევებსაც. მაგალითად, ის თუ როგორ მოქმედებს სისტემა „ბიჭებმა წერს“ წინადადების ანალიზის შემთხვევაში, ჩანს სურათ_51-ზე, სადაც სისტემამ გვითხრა რომ „ბიჭებმა“ ფორმა არასწორია და სწორი იქნება „ბიჭებს“ ფორმა.

სინტაქსური ანალიზი

ბიჭებმა წერს ანალიზი

წინადადება არასწორია

სიტყვა 'ბიჭებმა' ბრუნვა არ შეესაბამება ზმნას შესაძლო ფორმებია: ბიჭებს

სიტყვები:

- ბიჭებმა - არსებითი სახელი Nებმა ბრუნვაში
- წერს - ზმნა

სინტაქსური ფორმულა:

Nებმა [V [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]]

- [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]

სურათი 51 - სინტაქსური ანალიზი მოქმედებაში (4)

ასევე, სისტემა ასკვნის ნაცვალსახელი არის თუ არა სწორ ფორმაში. მაგალითად ვნახოთ რას გვეტყვის სისტემა წინადადებაზე „ისინი წერს“ (იხ. სურათი_52). - აქ სისტემა მიხვდა რომ სიტყვა „ისინი“ არის ნაცვალსახელი. ამასთან, წინადადებაზე გვითხრა, რომ არასწორია, რადგან ნაცვალსახელი ვერ ეთანხმება ზმნას და ასევე შემოგვთავაზა შესაძლო სწორი ფორმა და გვითხრა რომ ეს შეიძლება იყოს ნაცვალსახელი „ის“.

სინტაქსური ანალიზი

ანალიზი

წინადადება არასწორია

ნაცვალსახელი ვერ ეთანხმება ზმნას, შესაძლო სწორი ფორმაა 'ის'

სიტყვები:

1. ისინი - ნაცვალსახელი
2. წერს - ზმნა

სინტაქსური ფორმულა:

D [V [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]]

1. [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]

სურათი 52 - სინტაქსური ანალიზი მოქმედებაში (5)

ახლა განვიხილოთ წინადადება „ისინი გუშინ წერს“ (იხ. სურათი_53). აქ სისტემამ შეამჩნია, როგორც ის, რომ ნაცვალსახელი „ისინი“ არ ეთანხმება ზმნას, ასევე ისიც, რომ სიტყვა „გუშინ“, რომელიც გამოიღწო როგორც ზმნიზედა, ზმნას დროში არ უთანხმდება.

სინტაქსური ანალიზი

ანალიზი

წინადადება არასწორია

ნაცვალსახელი ვერ ეთანხმება ზმნას, შესაძლო სწორი ფორმაა 'ის' ზმნიზედა 'გუშინ' ზმნას დროში ვერ ეთანხმება

სიტყვები:

1. ისინი - ნაცვალსახელი
2. გუშინ - ზმნიზედა
3. წერს - ზმნა

სინტაქსური ფორმულა:

D Adverb [V [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]]

1. [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]

სურათი 53 - სინტაქსური ანალიზი მოქმედებაში (6)

ასევე, თუ სისტემამ წინადადებაში ზმნა ვერ მოძებნა ავტომატურად წერს რომ წინადადება არასწორია (იხ. სურათი_54).

სინტაქსური ანალიზი

ანალიზი

წინადადება არასწორია

წინადადებაში ზმნა არარის

სიტყვები:

1. ის - ნაცვალსახელი
2. დღეს - ზმნიზედა
3. წერილს - არსებითი სახელი Nს ბრუნვაში
4. ბავშვებს - არსებითი სახელი Nებს ბრუნვაში

სინტაქსური ფორმულა:
D Adverb Nს Nებს

სურათი 54 - სინტაქსური ანალიზი მოქმედებაში (7)

ასევე, სურათ_55-ზეა ნაჩვენებია, რომ სისტემამ, გაანალიზა რა წინადადება „ის დღეს წერილს წერს ბავშვებს“, დაასკვნა მისი სისწორე.

სინტაქსური ანალიზი

ანალიზი

წინადადება სწორია

სიტყვები:

1. ის - ნაცვალსახელი
2. დღეს - ზმნიზედა
3. წერილს - არსებითი სახელი Nს ბრუნვაში
4. წერს - ზმნა
5. ბავშვებს - არსებითი სახელი Nებს ბრუნვაში

სინტაქსური ფორმულა:
D Adverb Nს [V [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]] Nებს

1. [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]
2. ის [მას/Nს|მათ/Nებს] [Nს/Nებს]
3. ის წერილს [Nს/Nებს]
4. ის წერილს ბავშვებს

სურათი 55 - სინტაქსური ანალიზი მოქმედებაში (8)

ასევე, სურათ_56-ზე და სურათ_57-ზე ნაჩვენებია სისტემის მუშაობის შედეგები წინადადებებზე „ის გუშინწინ წერილს წერს ბავშვებს“ და „ის დღეს წერილს წერენ ბავშვებს“.

სინტაქსური ანალიზი

ის გუშინწინ წერილს წერს ბავშვებს

ანალიზი

წინადადება არასწორია

ზმნიზედა 'გუშინწინ' ზმნას დროში ვერ ეთანხმება

სიტყვები:

1. ის - ნაცვალსახელი
2. გუშინწინ - ზმნიზედა
3. წერილს - არსებითი სახელი Ns ბრუნვაში
4. წერს - ზმნა
5. ბავშვებს - არსებითი სახელი Nებს ბრუნვაში

სინტაქსური ფორმულა:

D Adverb Ns [V [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]] Nებს

1. [ის/Nი] [მას/Nს|მათ/Nებს] [Nს/Nებს]
2. ის [მას/Nს|მათ/Nებს] [Nს/Nებს]
3. ის წერილს [Nს/Nებს]
4. ის წერილს ბავშვებს

სურათი 56 - სინტაქსური ანალიზი მოქმედებაში (9)

სინტაქსური ანალიზი

ის დღეს წერილს წერენ ბავშვებს

ანალიზი

წინადადება არასწორია

ნაცვალსახელი ვერ ეთანხმება ზმნას, შესაძლო სწორი ფორმაა 'ისინი'

სიტყვები:

1. ის - ნაცვალსახელი
2. დღეს - ზმნიზედა
3. წერილს - არსებითი სახელი Ns ბრუნვაში
4. წერენ - ზმნა
5. ბავშვებს - არსებითი სახელი Nებს ბრუნვაში

სინტაქსური ფორმულა:

D Adverb Ns [V [ისინი/Nები] [მას|მათ / Nს]Nებს] [Nს/Nებს]] Nებს

1. [ისინი/Nები] [მას|მათ / Nს]Nებს] [Nს/Nებს]
2. [ისინი/Nები] [მასწერილსNებს] [Nს/Nებს]
3. [ისინი/Nები] ბავშვებს [Nს/Nებს]

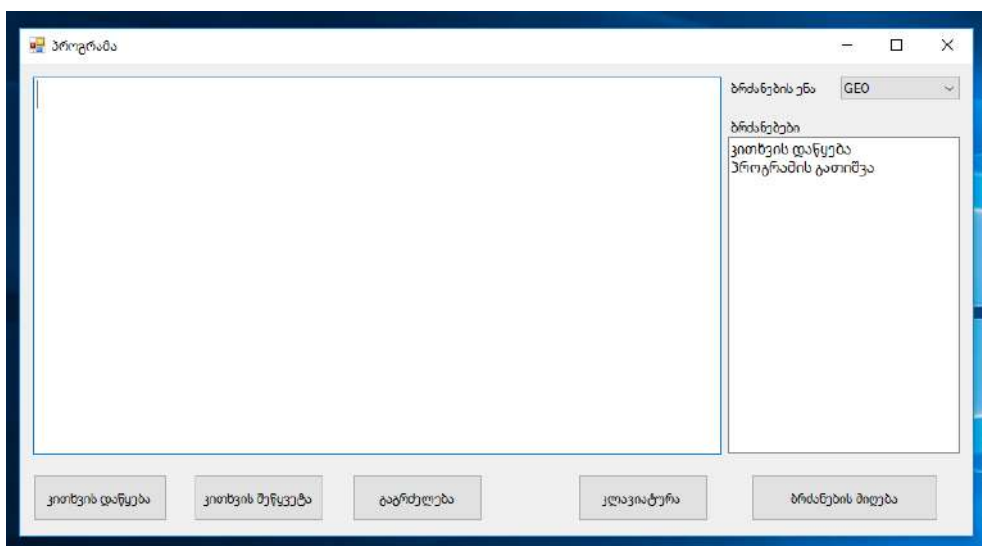
სურათი 57 - სინტაქსური ანალიზი მოქმედებაში (10)

2.10. ქართული ხმოვანმართვიანი მკითხველის გაფართოების გზით შემუშავებული ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები

ფხაკაძე-ჩიჩუას მეთოდებზე დაყრდნობით აიგო პროგრამა, რომელიც მოიცავს ქართული და ძველი ქართული ენებისა და აფხაზური ენის ხმოვანმართვიან მკითხველებს, რომლებიც აღჭურვილია ფუნქციებით „კითხვის დაწყება“, „კითხვის შეწყვეტა“, „კითხვის გაგრძელება“, რაც ხმოვანი და ლილაკიანი ბრძანებებით საშუალებას იძლევა ჩვენს მიერ აკრეფილი ტექსტი ჩავაყენოთ კითხვის რეჟიმში და მოვისმინოთ, სასურველ დროს გავაჩეროთ სინთეზირების ანუ მიმდინარე კითხვის პროცესი და შემდგომ, სურვილის შემთხვევაში, შეწყვეტილი ადგილიდან გავაგრძელოთ შეწყვეტილი კითხვის პროცესი. ამ მიზნით ჯერ შემუშავდა აფხაზური მეტყველების ამოცნობისა და ხმოვანი მართვის პირველი საცდელი ვერსიები. მოკლედ მათ შესახებ.

2.10.1. თანამედროვე აფხაზური მეტყველების ამოცნობისა და ხმოვანი მართვის პირველი საცდელი ვერსიები

სისტემის საწყისი იერსახე, რომელიც შესაბამისი ქართული სისტემების საფუძველზე აიგო, მოცემულია სურათ_58-ზე.



სურათი 58 - ქართულ-აფხაზური ხმოვანი მართვა

პროგრამის არეალი გაყოფილია სამ მნიშვნელოვან სივრცეთ. პირველი და ყველაზე დიდი ადგილი ეთმობა ტექსტურ ველს, რომელიც მოთავსებულია ზედა მარცხენა სივრცეში და დათმობილი აქვს მას მნიშვნელოვნად დიდი ადგილი. ეს არე განკუთვნილია ჩვენთვის სასურველი ტექსტის ჩასაწერად. პროგრამის ქვედა სივრცე დათმობილი აქვს ფუნქციურ დილაკებს. როგორც ზემოთაც აღინიშნა, ეს დილაკებია:

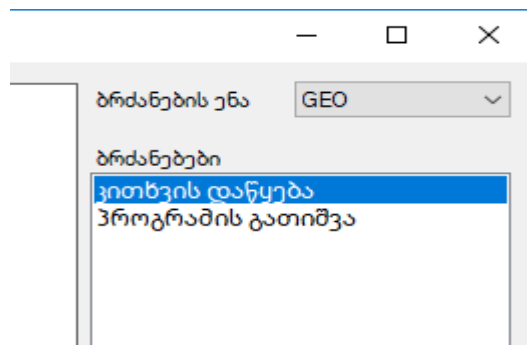
1. „კითხვის დაწყება“ დილაკი - ეს დილაკი უზრუნველყოფს ტექსტურ სივრცეში ჩაწერილი ტექსტის სინთეზირების ანუ კითხვის დაწყებას;

2. „კითხვის შეწყვეტა“ დილაკი. - ეს დილაკი უზრუნველყოფს დილაკი „კითხვის დაწყების“ შემდგომ ჩვენთვის სასურველ დროს მოვახდინოთ სინთეზირების ანუ კითხვის პროცესის შეჩერება;

3. „გაგრძელება“ დილაკი. - ეს დილაკი უზრუნველყოფს უკვე შეწყვეტილი კითხვის გაგრძელებას შეწყვეტილი ადგილიდან.

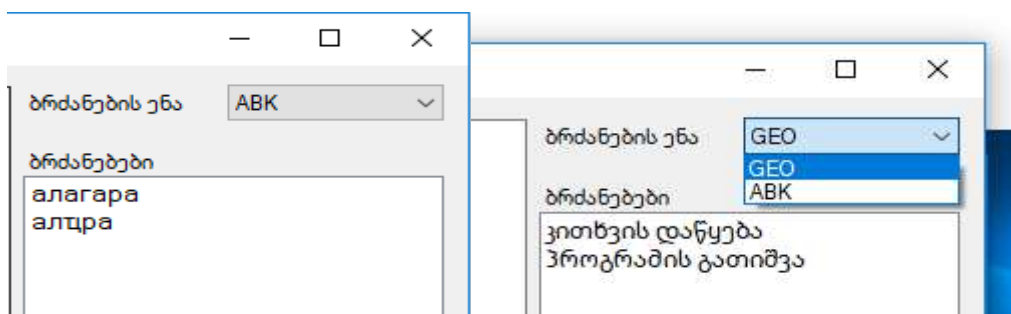
ასევე გვხვდება „კლავიატურა“ დილაკი, რომელიც უზრუნველყოფს გამოაჩინოს ფორმა სადაც დატანილი იქნება შემუშავებული დინამიური კლავიატურის ფორმა, რომელშიც იქნება ყველა ის სიმბოლო რომელიც ენაში გვხვდება და ასევე ბეჭდისთვის ფუნქციური დილაკები, როგორცაა ბოლო სიმბოლოს წაშლა და ჰარის (ანუ ცარიელი სივრცის) დაბეჭდვა. ასევე, აქ არის „ბრძანების მიღება“ დილაკი, რომლის გააქტიურების შემთხვევაში პროგრამა გადადის მოსმენის რეჟიმში და იწყება მითითებული ბრძანებებიდან მიცემულის გამოცნობა რეალურ დროში. „ბრძანების მიღება“ დილაკის გააქტიურებისას ჩნდება დილაკი „ბრძანების დასრულება“, რაც საშუალებას აძლევს მომხმარებელს რომ გაათავისუფლოს ანუ გამორთოს პროგრამა ხმოვანი ბრძანების გამოცნობის რეჟიმისგან, ანუ, ამ უკანასკნელის გააქტიურების შემთხვევაში პროგრამა გამოდის ხმოვანი ბრძანებების გამოცნობის რეჟიმიდან. გარდა ამისა, პროგრამაში ზედა მარჯვენა სივრცეში არის სექტორი, სადაც განთავსებულია ბრძანებების ენის ასარჩევი გრაფა და ასევე ბრძანებების სია, რომელიც ასახავს გამოსაცნობ ბრძანებების სიას სისტემის ფუნქციონირების მიმდინარე მომენტისათვის. „ბრძანების მიღება“

ლილაკზე დაჭერით პროგრამა ითვისებს ჩამონათვალში არსებულ ბრძანებების სიას და ცდილობს მიცემული ბრძანებებიდან გამოიყოს რომელი მათთაგანი იყო გახმოვანებული, რის შემდეგაც ასრულებს გამოცნობილი ხმოვანი ბრძანების შესაბამის ოპერაციას. მაგალითად თუ ჩვენ ვეტიყვით პროგრამას „პროგრამის გათიშვა“, ის ამოიცნობს ამ ბრძანებას და დახურავს პროგრამას. ბრძანების ამოცნობის შემთხვევაში, როგორც ეს სურათ_59-ზეა, პროგრამა გამოცნობილ ბრძანებას აფეროვნებს ლურჯად.



სურათი 59 - ქართული ხმოვანი მართვის ბრძანებები

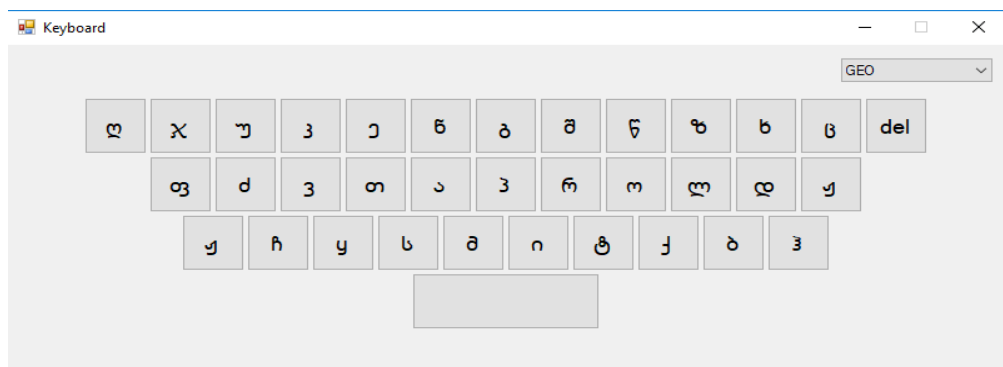
ბრძანების ენების სიაში მოცემული გვაქვს ორი ენა, ქართული და აფხაზური (იხ. სურათი_60). ენის შეცვლის შემთხვევაში ბრძანებების ჩამონათვალში გამოჩნდება არჩეულ ენაზე არსებული ბრძანებები. მაგალითად, ენის ლილაკის აფხაზურ ენაზე გადართვის შემთხვევაში, როგორც ეს სურათ_61-ზეა ნაჩვენები, ეკრანზე გაჩნდება აფხაზურ ბრძანებათა ის ჩამონათვალი, რომელიც ამ შემთხვევაში შინაარსობრივად ექვივალენტურია მანამდე მოცემული ქართული ბრძანებების სიის.



სურათი 60-61 - აფხაზური ხმოვანი მართვის ბრძანებები

2.10.2. ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები

დავუბრუნდეთ „კლავიატურა“ ლილაკზე დაჭერის ფუნქციურობას, მასზე დაჭერით ამოდის ფანჯარა სადაც დახატულია ლილაკების ერთობლიობა რაც ასახავს არჩეული ენის ანბანში არსებულ ყველა ასო სიმბოლოს და ასევე დართული აქვს ფუნქციური ლილაკები რომლებიცაა „del“ ლილაკი (მასზე დაჭერით ტექსტურ ველში იშლება ერთი სიმბოლო რომელიც კურსორიდან მარცხნივ მდგომია) და „space“ ლილაკი (მასზე დაჭერით ტექსტურ ველში ჩნდება ცარიელი სიმბოლო-ჰარი) ქვემოთ მოცემულია ქართული ენის კლავიატურის ნიმუში.



სურათი 62 - ქართული კლავიატურა (კომპვერსია)

კლავიატურის ფანჯარა საშუალებას გვაძლევს ლილაკის სიგრძე და სიგანე დავაკორექტიროთ ჩვენთვის სასურველად. ამასთან, კლავიატურა აწყობილია იმდაგვარად რომ საშუალებას გვაძლევს მარტივად შევცვალოთ ლილაკის მახასიათებლები, მაგალითად როგორიცაა სიგრძე, სიგანე, ასევე შევცვალოთ ლილაკების თანმიმდევრულობა, მაგალითად ქართული კლავიატურის სურათზე ასოები დალაგებულია იმ თანმიმდევრობით როგორი თანმიმდევრობითაცაა დატანილი ისინი ქართულ საბეჭდ მანქანაზე. ასევე შეგვიძლია განვსაზღვროთ როგორც „del“ ლილაკის, ისე „space“ ლილაკის ადგილები, ასოები დავაშოროთ ერთმანეთისგან ან მივაახლოვოთ ისინი უფრო ახლოს ვიდრე ეს სურათზეა აღნიშნული. კლავიატურა თავისით ითვლის სტრიქონში რამდენი ლილაკია და

ავტომატურად ასწორებს ჰორიზონტალური სწორებით ფორმის მარჯვენა და მარცხენა ბორტის მიმართ, რათა უფრო ლამაზად გამოჩნდეს ღილაკების განლაგება ვიზუალურად. ხაზგასასმელია ისიც, რომ კლავიატურის მოდელი კონფიგურირდება ძალიან მარტივად და იგი საშუალებას გვაძლევს შევქმნათ ნებისმიერი ენის კლავიატურა, ამ კუთხით იგი არანაირად არაა შეზღუდული, მათ შორის ამ მიდგომებით მარტივად შეიძლება აიგოს აგრეთვე ეროგლიფური დამწერლობის მქონე ენების კლავიატურებიც.

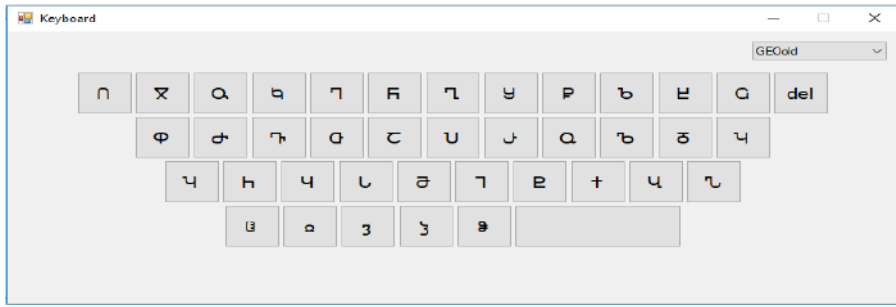
ქვემოთ სურათ_63-ზე მოცემულია ქართული კლავიატურის დასაკონფიგურირებელი მონაცემები:

```
new string[]
{
    "ღ ჯ უ კ ე ნ გ მ წ ზ ხ ც {del}",
    "ფ ძ ვ თ ა პ რ ო ლ დ ჟ",
    "ქ ჩ ყ ს მ ი ტ ქ ბ ჰ",
    "{space}"
}
```

სურათი 63 - ქართული კლავიატურის კონფიგურაციის მოდელი

ანუ, აქ საჭიროა მხოლოდ მივუთითოთ მწკრივზე რამდენი სიმბოლო განთავსდეს და სულ რამდენ ფენად იყოს დაყოფილი კლავიატურა. ქართული კლავიატურის მაგალითზე ჩვენ გვაქვს სულ ოთხი მწკრივი.

კლავიატურის ფანჯარის მარჯვენა ზედა კუთხეში გვხვდება მოსანიშნი, სადაც წერია ყველა ის ენა, რომლებზედაც მორგებულია კლავიატურა. ჩვენ შემთხვევაში ეს ენებია - ქართლი, ძველ ქართლი, და აფხაზური ენები. ამასთან, ენის გამოცვლის შემთხვევაში წაიშლება არსებული ღილაკების განლაგებები და ხელახლა მოხდება იმ ღილაკების გამოჩენა, რომლის ენაც ავირჩიეთ. მაგალითად, თუ ენად ძველ ქართულს ავირჩევთ, ჩვენი კლავიატურა ქვემოთ სურათ_64-ზე წარმოდგენილ სახეს.



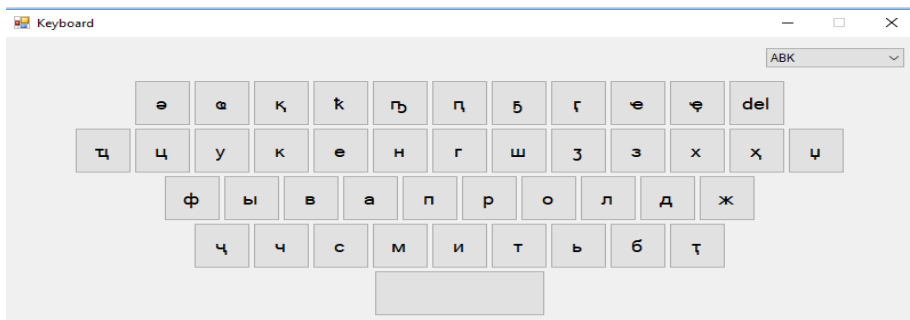
სურათი 64 - ძველი ქართული კლავიატურა (კომპ. ვერსია)

ახლა ჩვენ ვხედავთ თუ როგორ გამოიყურება ძველ ქართული კლავიატურა ჩვენს პროგრამაში და ქვემოთ, სურათ_65-ზე, წარმოდგენილია მისი დასაკონფიგურებელი მონაცემები.

```
new string[] {
    "ჩ x q w r f v p b l c {del}",
    "ფ ძ წ ც ს j q ზ ბ ყ",
    "ყ h ყ ს შ ნ e + ყ ლ",
    "შ o ვ ვ ჟ შ {space}"
}
```

სურათი 65 - ძველი ქართული კლავიატურის კონფიგურაცია

ანუ, როგორც ეს სურათიდანაც ჩანს, პროგრამა საშუალებას იძლევა, რომ კლავიატურის ფუნქციური ღილაკები, როგორცაა “del” და “space”, დავსვათ პანელის ნებისმიერ ადგილზე. ქვემოთ, სურათ_66-ზე ნაჩვენებია აფხაზურენოვანი კლავიატურა:



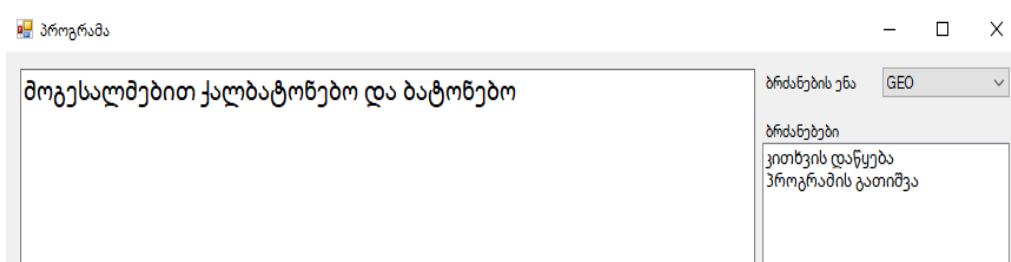
სურათი 66 - აფხაზური კლავიატურის კომპ. ვერსია

ამ შემთხვევაში კლავიატურაში ჩამატებულია ერთი სტრიქონი, რასთანაც თავისუფლად და დინამიურად ადაპტირდება არსებული კლავიატურის მოდელი. იგი აგებულია სურათ_67-ზე დატანილი საკლავიატურო კონფიგურაციის მიხედვით

```
new string[] {
    "ა ბ კ ლ მ ნ ი რ ე ფ {del}",
    "ც ც უ კ ე ნ გ შ ჯ ჰ ზ ხ ყ",
    "ფ ყ ვ ა პ რ ო ღ ჯ",
    "ყ ჩ ს მ ი თ ბ ტ",
    "{space}"
}
```

სურათი 67 - აფხაზური კლავიატური კონფიგურაცია

კლავიატურაზე აკრეფილი ტექსტი იწერება მთავარი პროგრამის ზედა მარცხენა სივრცეში განთავსებულ ტექსტურ ველში, ტექსტის ჩაწერის შემდგომ საშუალება გვეძლევა მოვისმინოთ ამ ტექსტის სინთეზირებული ანუ წაკითხული ვერსია შესაბამის ენაში როგორც „კითხვის დაწყება“ ღილაკზე დაჭერით ასევე შესაბამისი ხმოვანი ბრძანების მიცემითაც.

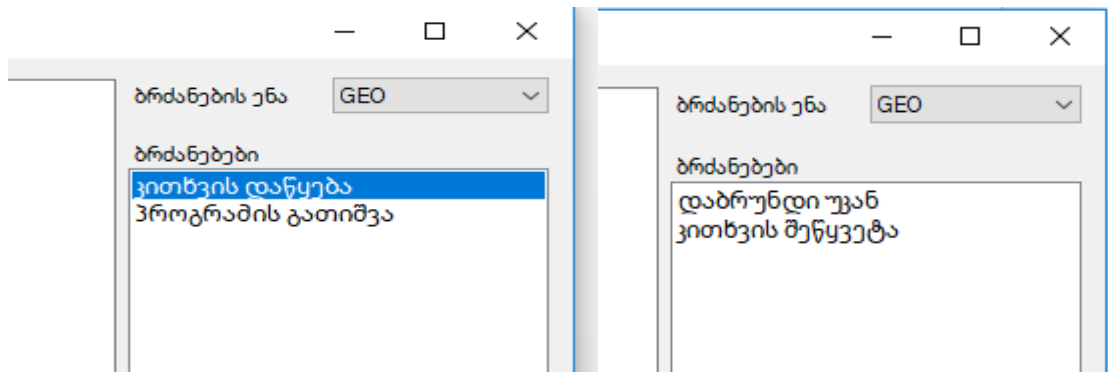


სურათი 68 - ქართული მკითხველის კომპ. ვერსია

ანუ, მთავარ ფორმაში (სურათი_68) მაღლა მარჯვენა სივრცეში მოცემული გვაქვს ქართულ ენაზე არსებული მიმდინარე ბრძანებების ჩამონათვალი. თუ დავაჭერთ „ბრძანების მიღება“ ღილაკს, რითაც გავააქტიურებთ ბრძანების მოსმენის რეჟიმს, და შემდგომ მიკროფონს ვეტყვით ბრძანებას „კითხვის დაწყება“, პროგრამა ამოიცნობს ანუ გაიგონებს

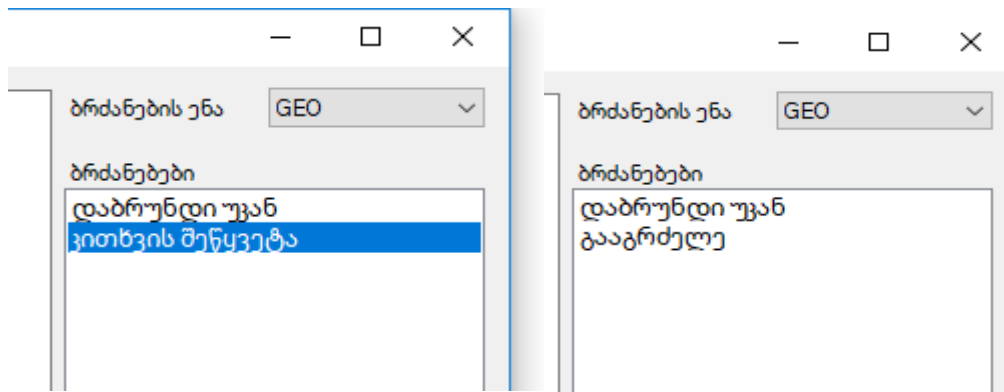
ანუ გაიგებს ამ ხმოვან ბრძანებას და გადავა მის შესრულებაზე ანუ დაიწყებს კითხვას, ანუ შეასრულებს იგივე ფუნქციონალს, რასაც აკეთებს „კითხვის დაწყება“ ლილაკის გააქტიურებისას.

ამასთან, კითხვის დაწყების პარალელურად ბრძანებების გრაფაში შეიცვლება ბრძანებების სია და ჩავარდება ის ბრძანებები, რომლებიც კორექტულია „კითხვის დაწყება“ ბრძანების შემდგომ ფუნქციონირებისთვის. კერძოდ, ეს ბრძანებებია „დაბრუნდი უკან“ და „კითხვის შეწყვეტა“ (იხ. სურათი_69) როგორცაა დატანილი სურათზე, მარცხნივ მოცემულია პირველი მდგომარეობა, ხოლო მარჯვნივ მოცემულია არჩეული ბრძანების შემდგომ ბრძანებების სიის შეცვლილი ანუ განახლებული ვერსია.



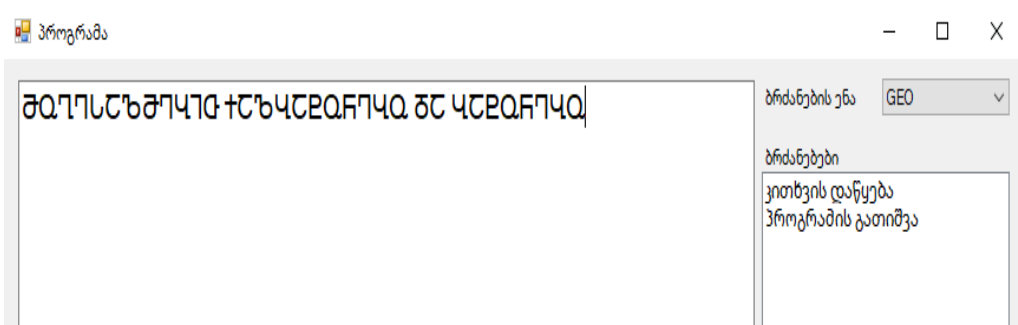
სურათი 69-70 - ქართული ხმოვანი ბრძანებების სტრუქტურა (1)

„დაბრუნდი უკან“ ბრძანების შემთხვევაში ბრძანებების სია გადადის საწყის გვერდზე ანუ საწყის ბრძანებებზე. „კითხვის შეწყვეტა“ ბრძანების შემთხვევაში ხდება იგივე რაც „კითხვის შეწყვეტა“ ლილაკზე დაჭერისას ანუ ჩვენ ტექსტურ გრაფაში არსებული ტექსტის სინთეზის პროცესი ანუ კითხვის პროცესი ჩერდება და ბრძანებების გრაფაში იყრება ბრძანებები „დაბრუნდი უკან“ და „გააგრძელე“ (იხ. სურათი_68).



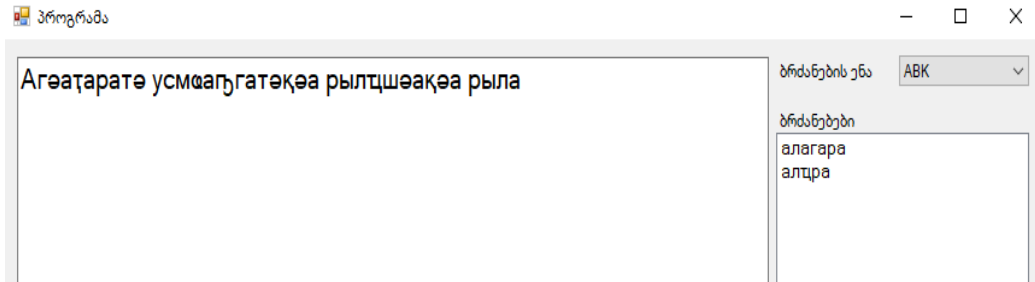
სურათი 71.1 - ქართული ხმოვანი ბრძანებების სტრუქტურა (2)

ამგვარად, როგორც ეს ასახულია სურათ_71.1, „დაბრუნდი უკან“ ბრძანების თქმის შემთხვევაში მდგომარეობა ბრუნდება პირვანდელ მდგომარეობაზე, ხოლო ბრძანება „გააგრძელე“ შემთხვევაში დაპაუზებული სინთეზირებული ხმა აგრძელებს ბოლო დაპაუზებული მდგომარეობიდან გაგრძელებას, ბრძანებების გრაფაში კი იწყება ბრძანებები როგორც სულ თავიდან „კითხვის დაწყება“ ბრძანების ამოცნობის დროს ხდებოდა.



სურათი 71.2 - ძველი ქართული მკითხველის კომპ. ვერსია

მკითხველი ქართული ხმოვანი ბრძანებებით ზუსტად ზემოთ აღწერილისნაირად მუშაობს ძველ ქართულ ტექსტებთანაც - ხმოვანი ბრძანებით იწყებს, აპაუზებს და აგრძელებს კითხვას (იხ. სურათი_71.2). ასევე მუშაობს პროგრამა აფხაზურ ხმოვან ბრძანებებზეც (იხ. სურათი_72). ანუ, ხმოვანი ბრძანებებით იწყება, პაუზდება და გრძელდება კითხვა.



სურათი 72 - აფხაზური მკითხველის კომპ. ვერსია

დაბოლოს, ქართულშიც და აფხაზურ ბრძანებებშიც მთავარი ბრძანებების სიაში არის „პროგრამის გათიშვა“ ბრძანებაც, რომლის ამოცნობის შემთხვევაშიც პროგრამა იხურება

2.11. ქართული N, A და V ტიპის სიტყვების ავტომატური მორფო-სინტაქსურ-სემანტიკური ანალიზატორის თითქმის სრული საცდელი და საცდელ-სამომხმარებლო ვერსიები

ქვემოთ მოკლედაა წარმოდგენილი ქართული N, A და V ტიპის სიტყვების ანალიზატორების ანუ ტაგერების, ლემატიზატორების, გენერატორებისა, დესკრიპტორებისა და ანოტატორების ის პირველი საცდელი და საცდელ-სამომხმარებლო საინტერნეტო ვერსიები და, აგრეთვე, მათი ის ახალი გაძლიერებული ვერსიები, რომელთა შემუშავება წინამდებარე სადისერტაციო კვლევის ერთი ძირითადი კვლევითი მიზანი იყო.

ფორმა1	ფორმა2	ტიპი (A/N)	სულიერი/უსული	რაოდენობა სულ									
ბავშვ	ბავშვ	N (არსებითი)	S (სულიერი)	115927									
ბავშვი	ბავშვიდან	ბავშვით	ბავშვის	ბავშვისთვის	ბავშვისკენ	ბავშვისგან	ბავშვს	ბავშვში	ბავშვზე	ბავშვად	ბავშვამდე	ბავშვთან	ბავშვმა
49471	279	674	29729	2492	165	168	22652	244	3466	329	61	2248	3949

სურათი 73 – N ანალიზატორის გაძლიერებული ვერსია (1)

სურათზე წარმოდგენილია N და A ტიპის სიტყვების ტაგერის (ანუ, ამომცნობის), დესკრიპტორისა (ანუ, აღმწერის), და გენერატორის (ანუ, მაწარმოებლის) საცდელ-სამომხმარებლო საინტერნეტო ვერსია. აქ სისტემამ ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და

მიდგომებზე დაყრდნობით გააანალიზა სიტყვა „ბავშვი“ და ამოიცნო ანუ მიხვდა იმას, რომ ის არის S კლასის N ტიპის ბრუნებადი სიტყვა ანუ სულიერი არსებითი სახელი. ასევე, სისტემამ ქართული ენის ლოგიკური გრამატიკის მიხედვით აწარმოა ანუ დააგენერირა მისი ყველა შესაძლო ლოგიკური ბრუნვითი ფორმა და ყველა ეს ბრუნვითი ფორმა დაადესკრიპტორა ანუ აღწერა როგორც მორფოლოგიური, ისე სინტაქსური თვალსაზრისით. ამასთან, აქვე ხაზს ვუსვამთ, რომ მორფოლოგიური აღწერა თავისთავად გულისხმობს აგრეთვე ამ სიტყვის ლემატიზაციას ანუ ამოცნობილი N ტიპის სიტყვაში ლემის ანუ ფუძის გამოყოფას [13].

სკამზე		filter	ანალიზი										
ფორმა1	ფორმა2	ტიპი (A/N)	სულიერი/უსული	რაოდენობა სულ									
სკამ	სკამ	N (არსებითი)	O (უსული)	23064									
სკამი	სკამიდან	სკამით	სკამის	სკამისთვის	სკამისკენ	სკამისგან	სკამს	სკამში	სკამზე	სკამად	სკამამდე	სკამთან	სკამმა
3586	2395	172	1452	84	290	4	1310	283	13210	11	57	190	20

სურათი 74 – N ანალიზატორის გაძლიერებული ვერსია (2)

ამ სურათზე სისტემამ ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით გააანალიზა სიტყვა „სკამზე“ და ამოიცნო ანუ მიხვდა იმას, რომ ის არის O კლასის N ტიპის ბრუნებადი სიტყვა ანუ უსულო არსებითი სახელი. ასევე, სისტემამ ქართული ენის ლოგიკური გრამატიკის მიხედვით აწარმოა ანუ დააგენერირა მისი ყველა შესაძლო ლოგიკური ბრუნვითი ფორმა და ყველა ეს ბრუნვითი ფორმა დაადესკრიპტორა ანუ აღწერა როგორც მორფოლოგიური, ისე სინტაქსური თვალსაზრისით [13].

დიდმა		filter	ანალიზი										
ფორმა1	ფორმა2	ტიპი (A/N)	სულიერი/უსული	რაოდენობა სულ									
დიდ	დიდ	A (ზედსართავი)	O (უსული)	464162									
დიდი	დიდიდან	დიდით	დიდის	დიდისთვის	დიდისკენ	დიდისგან	დიდს	დიდში	დიდზე	დიდად	დიდამდე	დიდთან	დიდმა
445087	22	24	1562	24	6	3	721	49	195	10855	5	53	5556

სურათი 75 - A ანალიზატორის გაძლიერებული ვერსია

აქ სისტემამ ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით გააანალიზა სიტყვა „დიდმა“

და გამოიცნო ანუ მიხვდა იმას, რომ ის არის A ტიპის ბრუნებადი სიტყვა ანუ ზედსართავი სახელი. ასევე, სისტემამ ქართული ენის ლოგიკური გრამატიკის მიხედვით აწარმოა ანუ დააგენერირა მისი ყველა შესაძლო ლოგიკური ბრუნვითი ფორმა და ყველა ეს ბრუნვითი ფორმა დაადესკრიპტორა ანუ აღწერა როგორც მორფოლოგიური, ისე სინტაქსური თვალსაზრისით [13].

წინსართი	ძირი	ფორმა	რაოდენობა	Z %	პირველი სვეტის	
წერ	ვ-ს_უღლება_2(2)		111545	100 %	0	
გწერ		312	გწერდი	61	გწერდე	10
წერ		793	წერდი	145	წერდე	32
მწერ		152	მწერდი	38	მწერდე	1
წერს		91813	წერდა	9723	წერდეს	89
გწერს		142	გწერდა	33	გწერდეს	3
მწერს		783	მწერდა	851	მწერდეს	11

სურათი 76 – V ანალიზატორის გაძლიერებული ვერსია (1)

აქ სისტემამ ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით გააანალიზა სიტყვა „წერ“ და:

1. მიხვდა იმას, რომ ის არის V ტიპის უღლებადი სიტყვა ანუ ზმნა;
2. მიხვდა იმას, რომ ის არის ორადგილიანი ვ-ს ტიპის უღლებით ანუ ვ-ს_2 უღლებით უღლებადი ზმნა;
3. მიხვდა იმას, რომ 'წერ' არის ამ ზმნის ძირი;
4. დაადესკრიპტორა ანუ აღწერა ის როგორც მორფოლოგიურად ისე სინტაქსურად;
5. დააგენერირა ანუ წარმოქმნა მისი მხოლოდითი რიცხვის უღლებითი ფორმები;
6. დააგენერირა ანუ წარმოქმნა მისი მხოლოდითი რიცხვის უღლებითი ფორმები როგორც -დი, ისე -დე მწკრივებში [13];

წინსართი	ძირი	ფორმა	რაოდენობა	Z %	პირველი სვეტის	
და	წერ	ვ-ს_ულლება_2(2)	3490	53 %	1	
	დაგწერ	0	დაგწერდი	0	დაგწერდე	0
	დაწერ	260	დაწერდი	31	დაწერდე	3
	დამწერ	3	დამწერდი	0	დამწერდე	0
	დაწერს	807	დაწერდა	350	დაწერდეს	0
	დაგწერს	7	დაგწერდა	0	დაგწერდეს	0
	დამწერს	59	დამწერდა	0	დამწერდეს	0

სურათი 77 - V ანალიზატორის გაძლიერებული ვერსია (2)

აქ ისევ და ისევ ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით სისტემამ ივარაუდა, რომ [13]:

1. „ვწერ“ ზმანსთან „და(...)" წინსართი არის 53%-იან ალბათურ თავსებადობაში;
2. ამდენივე პროცენტით ივარაუდა, რომ „დავწერ“ ზმნა, ისევე როგორც „ვწერ“ ზმნა არის ორადგილიანი ვ-ს ტიპის უღლებით ანუ ვ-ს_2(2) უღლებით უღლებადი ზმნა;
3. ქართული ენის ლოგიკური გრამატიკის მიხედვით დააგენერირა ანუ წარმოქმნა მისი სხვადასხვა შესაძლო უღლებითი ფორმები და ყველა ეს უღლებითი ფორმა დაადესკრიპტორა ანუ აღწერა როგორც მორფოლოგიური, სინტაქსური, დროითი და ასევე მათემატიკური ანუ პრედიკატული თვალსაზრისით (იხილეთ ქვემოთ წარმოდგენილი სურათები) [13].

წინსართი	ძირი	ფორმა	რაოდენობა	Z %	პირველი სვეტის	
გა	წერ	ვ-ს_ულლება_2(2)	121	48 %	0	
	გაგწერ	10	გაგწერდი	0	გაგწერდე	0
	გაწერ	6	გაწერდი	0	გაწერდე	0
	გამწერ	5	გამწერდი	0	გამწერდე	0
	გაწერს	80	გაწერდა	5	გაწერდეს	0
	გაგწერს	4	გაგწერდა	0	გაგწერდეს	0
	გამწერს	2	გამწერდა	1	გამწერდეს	1

სურათი 78 - V ანალიზატორის გაძლიერებული ვერსია (3)

და	წერ	ვწერ	t1 < t < t2	დაწერს(Sი,Sი,-)	მე Sს
და	წერ	გწერ	t1 < t < t2	დაწერს(Sი,Sი,-)	მე შენ
და	წერ	წერ	t1 < t < t2	დაწერს(Sი,Sი,-)	შენ Sს
და	წერ	მწერ	t1 < t < t2	დაწერს(Sი,Sი,-)	შენ მე
და	წერ	წერს	t1 < t < t2	დაწერს(Sი,Sი,-)	Sი Sს
და	წერ	გწერს	t1 < t < t2	დაწერს(Sი,Sი,-)	Sი შენ
და	წერ	მწერს	t1 < t < t2	დაწერს(Sი,Sი,-)	Sი მე
და	წერ	ვწერდი	t1 < t2 < t	დაწერს(Sი,Sი,-)	მე Sს
და	წერ	გწერდი	t1 < t2 < t	დაწერს(Sი,Sი,-)	მე შენ
და	წერ	წერდი	t1 < t2 < t	დაწერს(Sი,Sი,-)	შენ Sს
და	წერ	მწერდი	t1 < t2 < t	დაწერს(Sი,Sი,-)	შენ მე
და	წერ	წერდა	t1 < t2 < t	დაწერს(Sი,Sი,-)	Sი Sს
და	წერ	გწერდა	t1 < t2 < t	დაწერს(Sი,Sი,-)	Sი შენ
და	წერ	მწერდა	t1 < t2 < t	დაწერს(Sი,Sი,-)	Sი მე
და	წერ	დავწერ	t < tr	დაწერს(Sი,Sი,-)	მე Sს
და	წერ	დაგწერ	t < tr	დაწერს(Sი,Sი,-)	მე შენ
და	წერ	დაწერ	t < tr	დაწერს(Sი,Sი,-)	შენ Sს
და	წერ	დამწერ	t < tr	დაწერს(Sი,Sი,-)	შენ მე
და	წერ	დაწერს	t < tr	დაწერს(Sი,Sი,-)	Sი Sს
და	წერ	დაგწერს	t < tr	დაწერს(Sი,Sი,-)	Sი შენ
და	წერ	დამწერს	t < tr	დაწერს(Sი,Sი,-)	Sი მე

სურათი 79 - V ანალიზატორის გაძლიერებული ვერსია (4)

აქ ისევ და ისევ ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით სისტემამ ივარაუდა, რომ [13]:

1. „ვწერ“ ზმანსთან „გა(...)" წინსართი არის 24%-იან ალბათურ თავსებადობაში;
2. ამდენივე პროცენტით ივარაუდა, რომ „გავწერ“ ზმნა, ისევე როგორც „გავწერ“ ზმნის „ვწერ“ ვერსია არის ერთადგილიანი ვ-ს ტიპის უღლებით ანუ ვ-ს_2(2) უღლებით უღლებადი ზმნა;
3. ქართული ენის ლოგიკური გრამატიკის მიხედვით დააგენერირა ანუ წარმოქმნა მისი სხვადასხვა შესაძლო უღლებითი ფორმები და ყველა ეს უღლებითი ფორმა დაადესკრიპტორა ანუ აღწერა მორფოლოგიური, სინტაქსური, დროითი და ასევე მათემატიკური ანუ პრედიკატული თვალსაზრისით (იხ., ქვემოთ, სურათი 80) [13].

ჩა	წერ	ვ-ს_უღლება_2(2)	600	62 %	0	
	ჩაგწერ	20	ჩაგწერდი	1	ჩაგწერდე	0
	ჩაწერ	43	ჩაწერდი	4	ჩაწერდე	0
	ჩამწერ	213	ჩამწერდი	1	ჩამწერდე	0
	ჩაწერს	121	ჩაწერდა	55	ჩაწერდეს	0
	ჩაგწერს	5	ჩაგწერდა	0	ჩაგწერდეს	0
	ჩამწერს	23	ჩამწერდა	4	ჩამწერდეს	0

სურათი 80 – V ანალიზატორის გაძლიერებული ვერსია (6)

გა	წერ	ვწერ	$t1 < t < t2$	გაწერს(წი,წი,-)	მე წს
გა	წერ	გწერ	$t1 < t < t2$	გაწერს(წი,წი,-)	მე მენ
გა	წერ	წერ	$t1 < t < t2$	გაწერს(წი,წი,-)	შენ წს
გა	წერ	მწერ	$t1 < t < t2$	გაწერს(წი,წი,-)	შენ მე
გა	წერ	წერს	$t1 < t < t2$	გაწერს(წი,წი,-)	წი წს
გა	წერ	გწერს	$t1 < t < t2$	გაწერს(წი,წი,-)	წი მენ
გა	წერ	მწერს	$t1 < t < t2$	გაწერს(წი,წი,-)	წი მე
გა	წერ	ვწერდი	$t1 < t2 < t$	გაწერს(წი,წი,-)	მე წს
გა	წერ	გწერდი	$t1 < t2 < t$	გაწერს(წი,წი,-)	მე მენ
გა	წერ	წერდი	$t1 < t2 < t$	გაწერს(წი,წი,-)	შენ წს
გა	წერ	მწერდი	$t1 < t2 < t$	გაწერს(წი,წი,-)	შენ მე
გა	წერ	წერდა	$t1 < t2 < t$	გაწერს(წი,წი,-)	წი წს
გა	წერ	გწერდა	$t1 < t2 < t$	გაწერს(წი,წი,-)	წი მენ
გა	წერ	მწერდა	$t1 < t2 < t$	გაწერს(წი,წი,-)	წი მე
გა	წერ	გავწერ	$t < tr$	გაწერს(წი,წი,-)	მე წს
გა	წერ	გავწერ	$t < tr$	გაწერს(წი,წი,-)	მე მენ
გა	წერ	გაწერ	$t < tr$	გაწერს(წი,წი,-)	შენ წს
გა	წერ	გამწერ	$t < tr$	გაწერს(წი,წი,-)	შენ მე
გა	წერ	გაწერს	$t < tr$	გაწერს(წი,წი,-)	წი წს
გა	წერ	გავწერს	$t < tr$	გაწერს(წი,წი,-)	წი მენ
გა	წერ	გამწერს	$t < tr$	გაწერს(წი,წი,-)	წი მე

სურათი 81 – V ანალიზატორის გაძლიერებული ვერსია (5)

აქ ისევ და ისევ ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით სისტემამ [13]:

1. ივარაუდა, „ვწერ“ ზმანსთან „ჩა(...)" წინსართის 62%-იანი ალბათური თავსებადობა;
2. ამდენივე პროცენტით ივარაუდა, რომ „ჩავწერ“ ზმნა, ისევე როგორც „ჩავწერ“ ზმნის „ვწერ“ ვერსია არის ორადგილიანი ვ-ს ტიპის უღლებით ანუ ვ-ს_2(2) უღლებით უღლებადი ზმნა;
3. ქართული ენის ლოგიკური გრამატიკის მიხედვით დააგენერირა ანუ წარმოქმნა მისი სხვადასხვა შესაძლო უღლებითი ფორმები და ყველა ეს უღლებითი ფორმა დაადესკრიპტორა ანუ აღწერა მორფოლოგიური, სინტაქსური, დროითი და ასევე მათემატიკური ანუ პრედიკატული თვალსაზრისით. (იხ., ქვემოთ, სურათი_82).

ზემოთ ჩვენ ვრცლად, თუმცა მხოლოდ ნაწილობრივ წარმოგიდგინეთ ქართული N და A ტიპისა და V ტიპის სიტყვების ტაგერების, დესკრიპტორებისა, და გენერატორების საცდელ-სამომხმარებლო საინტერნეტო ვერსიები. ანუ, ჩვენ უკვე ნაწილობრივ ვნახეთ ამ

კომპლექსური სისტემების კომპლექსური ფუნქციონირების დეტალები. - ხაზგასმის მიზნით, კიდევ ერთხელ ვიმეორებთ, რომ სისტემამ ყველაფერი ის, რაც ზემოთ უკვე ვნახეთ, გააკეთა ავტომატურად ანუ ადამიანის ჩაურევლად ჩვენს მიერვე შემუშავებული ქართული ენის ვებ-კორპუსზე დაყრდნობით [13].

ჩა	წერ	ვწერ	$t1 < t < t2$	ჩაწერს(სი,სი,-)	მე სს
ჩა	წერ	გწერ	$t1 < t < t2$	ჩაწერს(სი,სი,-)	მე შენ
ჩა	წერ	წერ	$t1 < t < t2$	ჩაწერს(სი,სი,-)	შენ სს
ჩა	წერ	მწერ	$t1 < t < t2$	ჩაწერს(სი,სი,-)	შენ მე
ჩა	წერ	წერს	$t1 < t < t2$	ჩაწერს(სი,სი,-)	სი სს
ჩა	წერ	გწერს	$t1 < t < t2$	ჩაწერს(სი,სი,-)	სი შენ
ჩა	წერ	მწერს	$t1 < t < t2$	ჩაწერს(სი,სი,-)	სი მე
ჩა	წერ	ვწერდი	$t1 < t2 < t$	ჩაწერს(სი,სი,-)	მე სს
ჩა	წერ	გწერდი	$t1 < t2 < t$	ჩაწერს(სი,სი,-)	მე შენ
ჩა	წერ	წერდი	$t1 < t2 < t$	ჩაწერს(სი,სი,-)	შენ სს
ჩა	წერ	მწერდი	$t1 < t2 < t$	ჩაწერს(სი,სი,-)	შენ მე
ჩა	წერ	წერდა	$t1 < t2 < t$	ჩაწერს(სი,სი,-)	სი სს
ჩა	წერ	გწერდა	$t1 < t2 < t$	ჩაწერს(სი,სი,-)	სი შენ
ჩა	წერ	მწერდა	$t1 < t2 < t$	ჩაწერს(სი,სი,-)	სი მე
ჩა	წერ	ჩაწერ	$t < tr$	ჩაწერს(სი,სი,-)	მე სს
ჩა	წერ	ჩაგწერ	$t < tr$	ჩაწერს(სი,სი,-)	მე შენ
ჩა	წერ	ჩაწერ	$t < tr$	ჩაწერს(სი,სი,-)	შენ სს
ჩა	წერ	ჩამწერ	$t < tr$	ჩაწერს(სი,სი,-)	შენ მე
ჩა	წერ	ჩაწერს	$t < tr$	ჩაწერს(სი,სი,-)	სი სს
ჩა	წერ	ჩაგწერს	$t < tr$	ჩაწერს(სი,სი,-)	სი შენ
ჩა	წერ	ჩამწერს	$t < tr$	ჩაწერს(სი,სი,-)	სი მე

სურათი 82 - V ანალიზატორის გაძლიერებული ვერსია (7)

ამასთან, ისევ ვიმეორებთ: ის, რომ ამ ტიპის სხვა ქართული სისტემა არ არსებობს, ეს ჩვენ ვიცით ზუსტად. თუმცა, ჩვენს ხელთ არსებული მონაცემებით, რაც, ცხადია, შემდგომი დაზუსტების პროცესშია, ასეთი კომპლექსური ანუ მრავალფუნქციური სისტემა არ არსებობს არც არცერთ სხვა ენაში [13]. გარდა ამისა, ჩვენ ზემოთ უკვე გვქონდა საუბარი ქართული ტექსტების ანალიზატორის ანუ, ფიგურალურად, ქართული ენობრივი კომპიუტერული ტვინის სხვადასხვა ფრაგმენტულ რეალიზაციებზე. ამ კონტექსტში, ხაზს ვუსვამთ, რომ ზემოთ მოკლედ მიმოხილული სისტემებს უკვე მნიშვნელოვან წილად აქვთ აკრეფილი ქართული ენობრივი აზროვნებისთვის სასაფუძვლოდ მნიშვნელოვანი ანალიზური

შესაძლებლობები, რაც გასაგებს ხდის რომ ისინი ქართული ენობრივი კომპიუტერული ტვინის აგების მიმართულებით გადადგმული სასაფუძვლოდ მნიშვნელოვანი ნაბიჯებია.

ამგვარად, ხაზს ვუსვამთ, რომ წინამდებარე სადისერტაციო კვლევის ფარგლებში, ერთი მხრივ, ხარისხობრივად გაუმჯობესდა ზემოწარმოდგენილი სისტემების მოქმედების ხარისხი ანუ შეიქმნა ზემოაღწერილი სისტემების ახალი - გაუმჯობესებული ვერსიები და, ასევე, შეიქმნა ქვემოთ წარმოდგენილი ქართული არსებითი და ზედსართავის სახელების ანუ ქართული ბრუნებადი სიტყვების ანუ ქართული N და A ტიპის სიტყვების ავტომატური ანოტატორების საინტერნეტო სისტემის (ავტორები: კ. ფხაკაძე, მ. ჩიქვინიძე) ასევე ახალი - გაუმჯობესებული საცდელ-სამომხმარებლო ვერსიები (ავტორები: კ. ფხაკაძე, მ. ჩიქვინიძე, შ. მალიძე) [13].

სიტყვა

ჭურში = მხოლოებითი რიცხვის არსებითი სახელი -ში ბრუნვაში
ჭურ = მხოლოებითის ფუძე
-ში = ბრუნვის ნიშანი
ჭურ-ში = უფრჩხილებო მორფოლოგიური ფორმა
(ჭურ)ში = ფრჩხილებიანი მორფოლოგიური ფორმა
სულიერი/უსულო = უსულო
კითხვა: ჭურში - რაში?
ზოგადი ტიპი = Nში

სურათი 83 – N ტიპის ანოტატორის გაძლიერებული ვერსია (1)

აქ ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით სისტემამ გაანალიზა რა სიტყვა „ჭურში“ თითქმის ამომწურავად დააანოტირა ანუ დაახასიათა იგი როგორც მორფოლოგიური, ისე სინტაქსური ასევე ზოგადსემანტიკური თვალსაზრისით.

სიტყვა ვაშლებმა	სიტყვა დიდს
ანალიზი	ანალიზი
ვაშლებმა = მრავლობითი რიცხვის არსებითი სახელი -მა ბრუნვაში	დიდს ზედსართავი სახელი ს ბრუნვაში
ვაშლებ = მრავლობითის ფუძე	დიდ = მხოლობითის ფუძე
-ებ = მრავლობითი რიცხვის ნიშანი	-ს = ბრუნვის ნიშანი
ვაშლ = მხოლობითის ფუძე	დიდ-ს = უფრჩხილებო მორფოლოგიური ფორმა
-მა = ბრუნვის ნიშანი	(დიდ)ს = ფრჩხილებიანი მორფოლოგიური ფორმა
ვაშლ-ებ-მა = უფრჩხილებო მორფოლოგიური ფორმა	სულიერი/უსული = უსული
((ვაშლ)ებმა = ფრჩხილებიანი მორფოლოგიური ფორმა	კითხვა: დიდს = როგორ?
სულიერი/უსული = უსული	ზოგადი ტიპი = Aს
ზოგადი/კერძო = ზოგადი	
კითხვა: ვაშლებმა = რამ?	
ზოგადი ტიპი = Nებმა	

სურათი 84-85 - N და A ტიპის ანოტატორების გაძლიერებული ვერსია

სურათ_84-ზე ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით სისტემამ გაანალიზა რა სიტყვა „ვაშლებმა“ თითქმის ამომწურავად დააანოტირა ანუ დაახასიათა იგი როგორც მორფოლოგიური, ისე სინტაქსური, ასევე ზოგადსემანტიკური თვალსაზრისით. სურათ_85-ზე ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით სისტემამ გაანალიზა რა სიტყვა „დიდს“ თითქმის ამომწურავად დააანოტირა ანუ დაახასიათა იგი როგორც მორფოლოგიური, ისე სინტაქსური, ასევე ზოგადსემანტიკური თვალსაზრისით.

სიტყვა ღამაშებმა
ანალიზი
ღამაშებმა = მრავლობითი რიცხვის გასახელებული ზედსართავი -მა ბრუნვაში
ღამაშებ = მრავლობითის ფუძე
-ებ = მრავლობითი რიცხვის ნიშანი
ღამაშ = მხოლობითის ფუძე
-მა = ბრუნვის ნიშანი
ღამაშ-ებ-მა = უფრჩხილებო მორფოლოგიური ფორმა
((ღამაშ)ებმა = ფრჩხილებიანი მორფოლოგიური ფორმა
სულიერი/უსული = უსული
ზოგადი/კერძო = ზოგადი
კითხვა: ღამაშებმა = რამ?
ზოგადი ტიპი = Aმ(ნებმა)

სურათი 86 - N ტიპის ანოტატორის გაძლიერებული ვერსია

სურათ_86-ზე ქართული ენის ლოგიკური გრამატიკით განვითარებულ ახალ ხედვებსა და მიდგომებზე დაყრდნობით სისტემამ გაანალიზა რა სიტყვა „ლამაზებმა“ თითქმის ამომწურავად დააანოტირა ანუ დაახასიათა იგი როგორც მორფოლოგიური, ისე სინტაქსური, ასევე ზოგადსემანტიკური თვალსაზრისით. ამასთან, კიდევ ერთხელ ხაზს ვუსვამთ, რომ იმ ყველაფერს, რაც ზემოთ უკვე ვნახეთ, ეს უნიკალური ანუ სხვა ანალოგების არ მქონე სისტემები აკეთებენ არა იმის ხარჯზე, რომ კორპუსის შემადგენელი სიტყვების ნაწილი ადამიანის მიერ არის უკვე ანოტირებული ანუ აღწერილი, არამედ ამის გარეშე - თავიდან ბოლომდე ავტომატურად ანუ მხოლოდ სისტემაში ჩადგმულ ლოგიკასა და ალგორითმებზე დაყრდნობით, რომლებიც, თავის მხრივ, ფხაკაძე-ჩიქვინიძის მეთოდებით ფხაკაძის ქართული ენის ლოგიკურ გრამატიკით ქართულისთვის შემუშავებულ ახალ ხედვებზე დაყრდნობით [13].

2.12. ქართული ხმოვანი ლექსიკონის ახალი გაძლიერებული ვერსია

კორპუსში სადისერტაციო კვლევის შედეგად ჩადგმული ქართული ხმოვანი ლექსიკონის ახალი გაძლიერებული ვერსია იყენებს მასში არსებულ ქართულიდან ინგლისურად და გერმანულად ნათარგმნ მონაცემებს.

ლექსიკონი

სიტყვა

ძებნა

მონაცემები

ID	Word	Translate	TranslateLanguage
1	ის	He	ENG
2	ის	Die	GER
3	მუშაობას	Work	ENG
4	მუშაობას	Arbeits	GER
5	მომწევიტა	Mometsqita	ENG
6	მომწევიტა	Mometsqite	GER
7	გამუდმებით	Continuously	ENG
8	გამუდმებით	Ständig	GER
9	ჩავჩინებდი	Chavchichinebdi	ENG
10	ჩავჩინებდი	Chavchichinebdi	GER

1 2 3 4 5 --

სურათი 87 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (1)

87-ე სურათზე ნაჩვენებია ლექსიკონის მონაცემები სადაც გვხვდება 60000 მდე ნათარგმნი ქართული სიტყვა როგორც ინგლისურად ასევე გერმანულად. სისტემა საშუალებას გვაძლევს ნათარგმნი მონაცემები დავაკორექტიროთ თუ მასში შევამჩნევთ ხარვეზს. მაგალითად მოვძებნოთ სიტყვა „კაცი“, ვიქცევით შემდეგნაირად „სიტყვა“ ველში ვწერთ სიტყვას „კაცი“ და ვაჭერთ ქვემოთ არსებულ ღილაკს „ძებნა“ [13].

The screenshot shows a search interface for the word 'კაცი' (man) in Georgian. It includes a search bar with the word entered, a 'ძებნა' (search) button, and a table of results. The table has columns for 'სიტყვა' (word), 'თარგმანი' (translation), and 'ენა' (language). The first row shows 'კაცი' translated to 'Male' in English (ENG) with a 'შენახვა' (save) button. The second row shows 'კაცი' translated to 'Männlich' in German (GER) with a 'შენახვა' (save) button.

სურათი 88 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (2)

სურათზე ასახულია მონაცემები რომელიც „კაცი“ სიტყვის ძებნის შემდეგ გამოგვიტანა სისტემამ, კერძოდ, მოგვცა ნათარგმნი ინგლისური და ქართული შესატყვისი რომლებიც მოცემულია რედაქტირებად ველში, რაც საშუალებას გვაძლევს სურვილისამებრ შევცვალოთ ნათარგმნი მასალა და დავაჭიროთ „შენახვა“ ღილაკს ხელი, რაც გამოიწვევს დაკორექტირებული მონაცემების შენახვას მონაცემთა სისტემაში. ქვემოთ მოცემული ლექსიკონის თავდაპირველი სახე გამოიყურება მინიმალისტურად, პანელზე თავიდან ჩანს მხოლოდ ტექსტური ველი და ღილაკი „ძებნა“ (იხ., ქვემოთ, სურათი_89) [13].

The screenshot shows a simplified search interface for the word 'კაცი' (man) in Georgian. It features a search bar with the word entered and a 'ძებნა' (search) button.

სურათი 89 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (3)

საძიებო ველში უნდა შევიტანოთ ჩვენთვის სასურველი სიტყვა და მოდით ეს იყოს სიტყვა „კაცი“ და შემდგომ დავაჭიროთ ძეზნა ღილაკს.

ლექსიკონი

სიტყვა



ინგლისური

Male

გერმანული

Männlich

ძეზნა

წინადადებების ჩვენება

კონტექსტის ჩვენება

სურათი 90 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (4)

როგორც სურათ_90-ზე ვხედავთ სიტყვა „კაცი“-სთვის სისტემამ დაგვიბრუნა მისი ინგლისური შესატყვისი და ასევე გერმანული შესატყვისი.

ასევე სისტემა საშუალებას გვაძლევს რომ სიტყვის მარჯვნივ განთავსებულ წითელ დაკვრის ღილაკზე დაჭერით მოვისმინოთ არსებული სიტყვისა თუ მისი ნათარგმნი შესატყვისის გახმოვანება იმ ენაზე რა ენაზეც ნათარგმნია სიტყვა. თუ დავაჭერთ ინგლისური ნათარგმნი სიტყვის მარჯვენა „დაკვრა“ ღილაკს მოვისმენთ ჩვენი ჩაწერილი სიტყვის ინგლისური ვერსიის გახმოვანებას ინგლისურად, ხოლო გერმანულის შემთხვევაში სიტყვა იქნება წარმოთქმული გერმანულად. ასევე, თუ მოძებნილი სიტყვის მარჯვნივ მდებარე წითელ ღილაკს დავაჭერთ, მაშინ მოვისმენთ ქართული სიტყვის გახმოვანებას ქართული მკითხველის მეშვეობით. ასევე სისტემამ მოგვცა ორი ღილაკი ესენია „წინადადებების ჩვენება“ და „კონტექსტების ჩვენება“. პირველ ღილაკზე დაჭერით

დავინახავთ რამდენიმე წინადადებას სადაც ჩვენი საძიებო სიტყვა იქნება ნახსენები [13].

წინადადების ჩვენება
ოქროშეილი მანძინგ კი, როდესაც უმეტესად იყო, მიეწერე და უფროსი, ახლა მანძინგ თქვი სინარტულეთო, მაშინ მეცნა, რომ ის მანძინგ კაცი იყო და მეცნა, ნაციონალური მოძრაობისგან განსხვავებით უფრო მეტი ვაგვიცხობა ჰქონდა. წყარო
ბევრი კაცი უზრალოდ გარბის საღვარელო ადამიანისაგან, როცა მას შესაბამისი პირობები არ გააჩნია. წყარო
ყველა კაცი ერთნაირია, მათ არ შეუძლიათ ერთგულება, დალატობენ ყველაზე ღირსეულ ქალებსაც, არ გააჩნიათ გრძნობები, არ აინტერესებთ რწმანტიკა და პირადელ ურთიერთობებში მხადრად არ გამოხატავენ ემოციებს, იმიტომ რომ ეს უმოციები საერთოდ არ გააჩნიათ - აი ხრალდებები, რომლებიც ყველაზე ხშირად ისმის მათი მისამართით, მაგრამ არის კი ეს სინარტულეთო? წყარო
მე პრინციპებს კაცი ვარ და აუცილებლად გავფრტვებ ბრძოლას, აუცილებლად დაეიხსნით იმ ურთი კაცისგან ქვეყანას, - განაცხადა გიორგი რურუამ. წყარო
შეიძლება უფროად გამოვიდეს გზაზე კაცი გადამხობის. წყარო
ის რომ რიგითი ხუნდაძე იყოს, კიდევ აიტანდა კაცი, მაგრამ კობახიძე ოცნების ერთ-ერთი ლიდერია, ისეთი დიდურა, რომელიც, მძიმე შედეგობების მიუხედავად, არ მიიცილეს და, მისი ტელეინტერვიუებით თუ ვიზუალურად, დაწინაურებასაც კი უპირობებდნენ სიის სათავეში. წყარო
რამდენიმე კაცი ცხოვრობს სულ აქ. წყარო
ისტორიკოსი კაცი ხარ და გვიყვინება იყო ხალხი, ვინც იარაღით ხელში იბრძოდა სამშობლოსათვის, გმირებად იქნენ, ხოლო ისინი, ვინც ამ ომს გაეტეა, ლაჩრებად მოინათლებიან. წყარო
დებილები ვართ ამა რა ვართ, ერთლარაინ წამალს, რომ ას ლარად ასალებს ჩვენზე, მამისასხლად რომ ყიდის, დებილები ვართ, ამა რა, სწორე გეითხობა კი, დებილები ვართ ასეთი კაცი რომ გვიხის ვარლამტენი და კიდევ ჭკუას რომ გვასწავლის და კიდევ დებილებს გუბახის. წყარო
რაც შეეხება ოპოზიციის, ნაციონალური მოძრაობის ოპტიმიზმს, ეს ოპტიმიზმი იყო კარგად გამოხატული გასულ დღეებში, ნახეთ, რომ მათ ვერ შეძლეს მხარდამტერების მობილიზება, 5 000 კაცი იდგა მათ მიერ ატვივის ხარჯზე მობილიზებულ ატვივზე, შესაბამისად, ეს უჩვენებს იმ სურათს, რომელიც აუცილებლად აისახება შემდეგ საპარლამენტო არჩევნებში, - განაცხადა კობახიძემ. წყარო

სურათი 91 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (5)

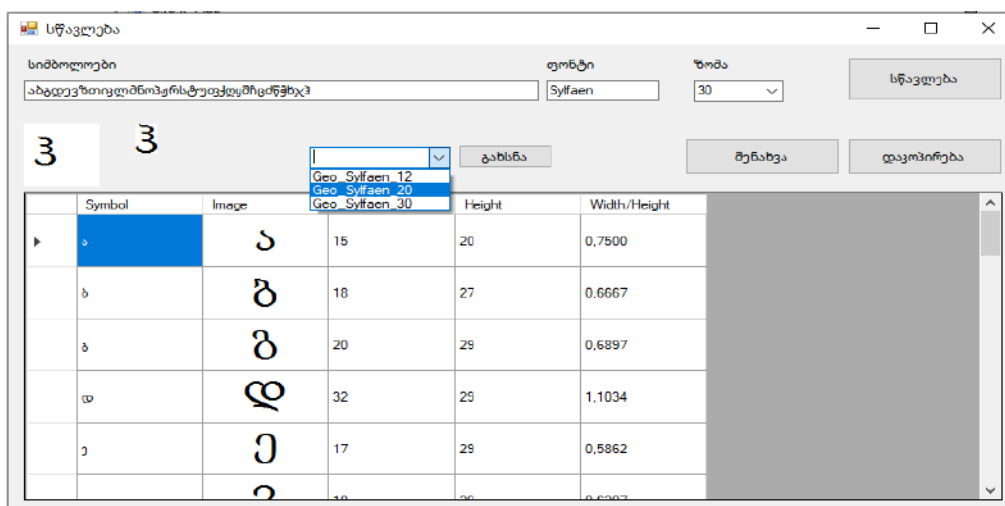
როგორც სურათზე ჩანს, გამოვიდა ის წინადადებები, სადაც ჩვენი მოძებნილი სიტყვაა და ასევე დაერთო იმ წყაროს ლინკი, რომლიდანაც ეს წინადადებაა. მეორე ღილაკზე დაჭერით გამოდის სიტყვის შემცველი კონტექსტები.

კონტექსტის ჩვენება
მანძინგ კაცი
გზაზე, კაცი
ასეთი კაცი
კაცი იყო
კაცი უზრალოდ
კაცი ერთნაირია
კაცი ვარ
კაცი ცხოვრობს
მანძინგ კაცი იყო
ასეთი კაცი რომ

სურათი 92 - ხმოვანი ლექსიკონის გაძლიერებული ვერსია (6)

კორპუსში არსებულ მონაცემებზე დაყრდნობით სურათზე ჩანს სიტყვა „კაცი“ რამდენიმე კონტექსტში

სწავლების პროცესში პროგრამა შეყვანილ სიმბოლოებს ჩამოუვლის და დახატავს არჩეულ შრიფტში არჩეული ზომით, მოაშორებს თეთრ სივრცეებს და აიღებს მხოლოდ იმ მართკუთხედს სურათიდან რასაც დასასწავლი სიმბოლო იკავებს (ეს პროცესი წარმოდგენილია სურათი_95-ით). შეყვანილი სიმბოლოების ქვემოთ გამოჩნდა თეთრი კუბიკი სადაც არის ჩახატული ერთ-ერთი დასასწავლი სიმბოლო, შემდგომ ამ სურათიდან სისტემამ ამოიღო მის მარჯვნივ მდგარი სურათი, სადაც მხოლოდ ის ფართობია, რასაც სიმბოლო იკავებს, ასევე დააფიქსირა სურათის სიგრძე და სიგანე პიქსელებში, ასევე სიგრძის და სიგანის თანაფარდობა. ასე მოიქცა ყველა სიმბოლოზე და სურათზე თქვენ ხედავთ სიას სადაც ჩაყრილია უკვე დასასწავლად გამზადებული სიმბოლოები თავისი ვიზუალური მონაცემებით. როცა პროცესი მორჩება და პროგრამა ყველა სიმბოლოს გაივლის, გვაქვს საშუალება, რომ ჩვენი დასწავლილი მონაცემები შევინახოთ. შესანახად საჭიროა დავაჭიროთ ღილაკს „შენახვა“, რის შემდგომაც გამოჩნდება დიალოგური ფანჯარა, რაც მოგვთხოვს შესანახ ფაილზე სახელის დარქმევას და სახელის მითითების შემდგომ შეგვეძლება ამ მონაცემების შენახვა.



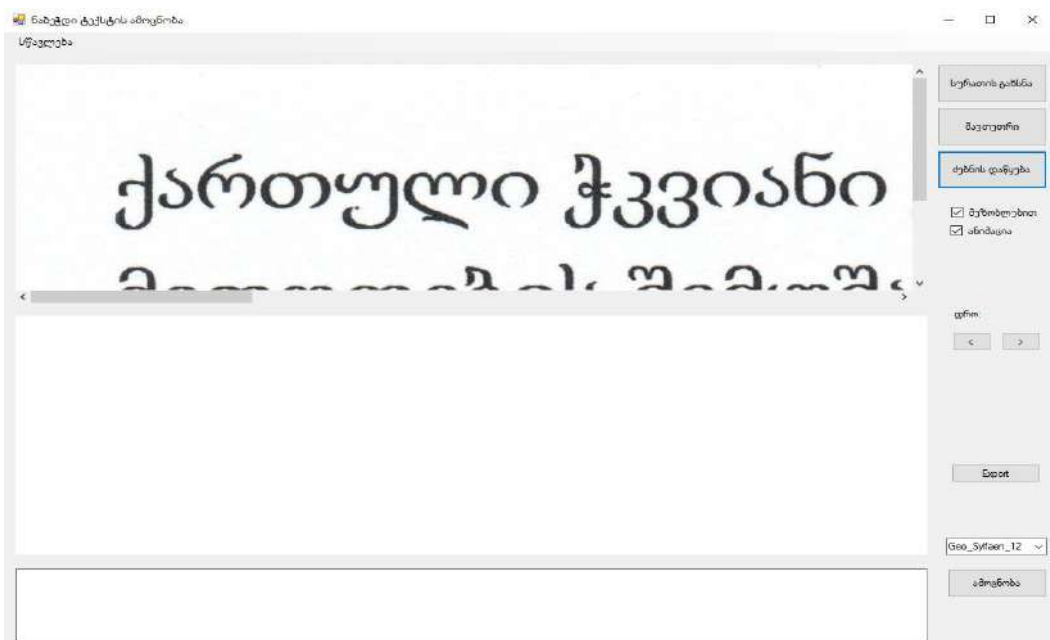
სურათი 95 - ქართული ტექსტის ამომცნობი (3)

ქვემოთ. სურათ_96-ზე, თქვენ ხედავთ რომ პროგრამაში უკვე არის შენახული რამდენიმე დასწავლილი მონაცემი. თუ ავირჩევთ სასურველს და

დავაჭერთ გახსნას, ქვედა გრაფაში გამოჩნდება თუ რა დასწავლილ მასალას მოიცავს არჩეული მონაცემები. ახლა ვისაუბროთ სურათზე სავარაუდო სიმბოლოების ფარგლების დადგენაზე. პროგრამას ვუთითებთ გასაანალიზებელ სურათს (იხ. სურათი_96), ძირითადად ტესტირება ხდება ხარისხიანად დასკანერებულ სურათზე, სადაც დატანილია ტექსტი სილფაინით აკრეფილი ზომა 12-ში, რაც ყველაზე ფართოთ გამოყენებადი ფონტია.

ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება

სურათი 96 - ამოსაცნობი ტექსტის სურათი



სურათი 97 - ქართული ტექსტის ამომცნობი (4)

სისტემის პირველი სახე (იხ. სურათი_97) შედგება ორი სასურათო სივრცისგან, პირველ ანუ ზედა სასურათო სივრცეში განთავსებულია ის მასალა, რაზეც პროგრამა მუშაობს ანუ ტექსტის ის სურათი, რაც უნდა გაანალიზოს. ქვედა სასურათო სივრცე არის იმისათვის, რომ პროგრამამ გაანალიზების თითოეულ ეტაპზე ვიზუალურად გვანახოს, თუ რას აკეთებს. თავდაპირველად ეს სივრცე არის თეთრი. პროგრამა მუშაობას იწყებს „ტექსტის დაწერა“ ლილაკის გააქტიურებით. ალგორითმი მუშაობს შემდეგნაირად

იწყებს ზევიდან ქვევით სურათის გაანალიზებას სიგანით და სიმაღლით ერთი პიქსელი, მარცხნიდან მარჯვნივ როცა ნახავს შავ პიქსელს შემდგომ ხდება სტრიქონის დადგენა. სტრიქონის დადგენა ხდება საწყისი შავი პიქსელით. მერე აქ უკვე ერთვება სიმბოლოს ჩარჩოების პოვნის ალგორითმი. ამისთვის ერთი შავი პიქსელით დავადგენთ ყველა ამ შავი პიქსელის მეზობლებს, მანამდე სანამ არ შეგვხვდება თეთრი პიქსელები. შედეგად ჩვენ გვაქვს პირველი სტრიქონის ყველაზე მაღალი სიმბოლოს საზღვრები. შემდგომ ემებს პროგრამა უკვე აღმოჩენილი საზღვრის მარცხნივ ან მარჯვნივ, ხომ არ არის შავი წერტილები. როცა იპოვის შავ წერტილს, ამოიღებს ამ წერტილით შემადგენელი სიმბოლოს საზღვრებს და ისევ ნახავს ამ სიმბოლოს მარცხნივ ან მარჯვნივ ხომ არარის შავი წერტილები. სტრიქონის ამოცნობის დასრულება ხდება მაშინ, როცა დარჩენილ მონაკვეთებში ვერ ნახავს შავ წერტილს. ამოცნობის დასრულების შემდგომ სურათიდან ამოშლის ნაპოვნ სიმბოლოებს (გათეთრებს უკვე ნაპოვნ ფართობებს) და ისევ დაიწყებს თავიდან ანუ გაიმეორებს ერთხელ უკვე გავლილ გზას და ა.შ. ციკლურად სურათის გასრულებამდე. შედეგად მეორე სასურათო სივრცეში ჩნდება პროგრამის მიერ გენერირებული სურათი (იხ. სურათი_98), რომელშიც ყველა სავარაუდო სიმბოლო მოქცეულია მართკუთხედში

ქართული ქვევანი კორპუსის ახალი განმავითარებული ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება



სურათი 98-99 - ქართული ტექსტის ამომცნობი (5)

კიდევ უფრო მეტი თვალსაჩინოებისათვის სურათ_100-ზე ჩანს, რომ საწყისი ტექსტის პირველი სტრიქონის ყველაზე მაღალ წერტილად სისტემამ მიიჩნია

სტრიქონში არსებული ბოლო „ს“ სიმბოლო. აქედან შეგვიძლია დავასკვნათ რომ დასკანერებული სურათი ოდნავ დახრილია რადგან სრული სისწორის შემთხვევაში უნდა აერჩია სტრიქონში არსებული მარცხნივ მდგომი „ს“ სიმბოლო. ამ პირველი ორი ეტაპის გავლის შემდეგ იწყება სიმბოლოს ამოცნობის პროცესი.

ResizedSource	Model	Compare	Same Pix	%
თ	თ	99	509	91.22
თ	ო	99	334	68.72
თ	ი	99	295	65.28
თ	უ	99	518	63.79
თ	ტ	99	606	65.66
თ	ც	99	387	63.55
თ	რ	99	438	62.54
თ	ე	99	306	62.07
თ	მ	99	300	61.79

Source	Resized/Model	Compare	Total Pix	Same Pix	NotCompare	%
თ	თ	99	3465	3149	316	90.88
თ	ო	99	3465	2628	807	75.64
თ	ი	99	3465	2301	1164	66.41
თ	უ	99	3465	2281	1184	65.63
თ	ტ	99	3465	2282	1203	65.28
თ	ც	99	3465	2250	1215	64.94
თ	რ	99	3465	2220	1245	64.07
თ	ე	99	3465	2215	1250	63.92
თ	მ	99	3465	2200	1265	63.49

სურათი 100 - ქართული ტექსტის გამომცნობი (6)

იმისათვის, რომ ვნახოთ როგორ მუშაობს სიმბოლოს ამოცნობის პროცესი საკმარისია სასურველ ასოს ორჯერ დავაჭიროთ „მაუსი“, შედეგად ამოდის ფანჯარა, სადაც ნათლად არის ასახული ამორჩეულ სურათს დასწავლილი მოდელი რა სავარაუდო სიმბოლოდ აღიქვავს.

მაგალითად განვიხილოთ პირველი სტრიქონის მეოთხე სურათი, რომელშიც მოთავსებულია „თ“ ასო, მაგრამ ეს ჯერ სისტემამ არ იცის. მასზე ორჯერ დაკლიკებით ამოვა შემდეგი ფანჯარა (იხ. სურათი_101).

მარცხნივ პროგრამა გვიწერს ამოსაცნობი სურათის მონაცემებს როგორც სიგრძის და სიგანის ასევე მისი თანაფართობის, მარჯვნივ კი მოცემულია ამოცნობის ორი მიდგომა.

ორივე მიდგომა ითვალისწინებს შემდეგს: დასადარებელიც და დამადარებელი სურათი უნდა იყოს ერთიდაიგივე რეზოლუციაში, რათა მოხდეს მათი დადარება. დამადარებელი არის ჩვენს შემთხვევაში მარცხნივ არსებული სურათი, სადაც ასახულია ასო ბგერა „თ“.

თ	თ	თ	თ
თ	ჲ	თ	თ
თ	ი	თ	თ
თ	ჳ	თ	თ
თ	ტ	თ	თ
თ	ც	თ	თ
თ	წ	თ	თ
თ	ე	თ	თ
თ	მ	თ	თ

სურათი 101-102 - ქართული ტექსტის ამომცნობი (7)

მარცხენა მეთოდში დამადარებელი გადადის დასადარებელ რეზოლუციაზე ანუ ჩვენს შემთხვევაში „თ“-ს სურათი გადავა 33-ვე ასოს რეზოლუციაზე და ისე დაედარება მათ (იხ. სურათი_101). მარჯვენა მეთოდში დამადარებელი არ იცვლის რეზოლუციას და იგი მოქმედებს ყველა დასადარებელის რეზოლუციაზე, ანუ 33-ვე ფორმა იცვლის სიგრძე სიგანეს და იღებს დამადარებელის სიგრძესა სიგანეს და შემდგომ ხდება დადარება (იხ. სურათი_102). ანუ, მოკლედ რომ ვთქვათ პირველ შემთხვევაში დამადარებელი იცვლის ფორმებს, ხოლო მეორე შემთხვევაში ფორმებს იცვლის დასადარებელი.

ეს ორი მიდგომა ცდებიდან გამომდინარე გვაძლევს სხვადასხვანაირ შედეგებს როგორ ხდება ორი სურათს შორის ალბათური თავსებადობის ამოღება. რადგან დამადარებელიც და დასადარებელი სურათებიც ერთიდაიგივე რეზოლუციაში გვაქვს უკვე შესაძლებელია რომ ისინი პიქსელის სიზუსტით დავადართო ერთმანეთს, ორი სურათის დადარების მეთოდი ორივე მიდგომაში ერთნაირია და მუშაობს შემდეგნაირად: ადარებს

მარცხენა სურათს მარჯვენას და აკეთებს მესამე სურათს სადაც ასახულია თუ რამდენია ისეთი წერტილი სადაც მოხდა თეთრი ფერების დამთხვევა რომელსაც ასახავს ყვითელ ფერად, თუ რამდენია ისეთი წერტილი სადაც მოხდა შავი ფერების დამთხვევა, რომელსაც ასახავს წითელ ფერად და თუ რამდენია ისეთი პიქსელები დასაც ვერ მოხდა თანხვედრა ასეთ პიქსელებს ასახავს მწვანე ფერებად, როგორც ეს ნაჩვენებია სურათ_103-ზე.

Source	ResizedModel	Compare	Total PX	Same PXs	NotCompare	%
			3465	3149	316	90,88
			3465	2628	837	75,84
			3465	2301	1164	66,41
			3465	2281	1184	65,83
			3465	2262	1203	65,28
			3465	2250	1215	64,94
			3465	2220	1245	64,07
			3465	2215	1250	63,92
			3465	2200	1265	63,49

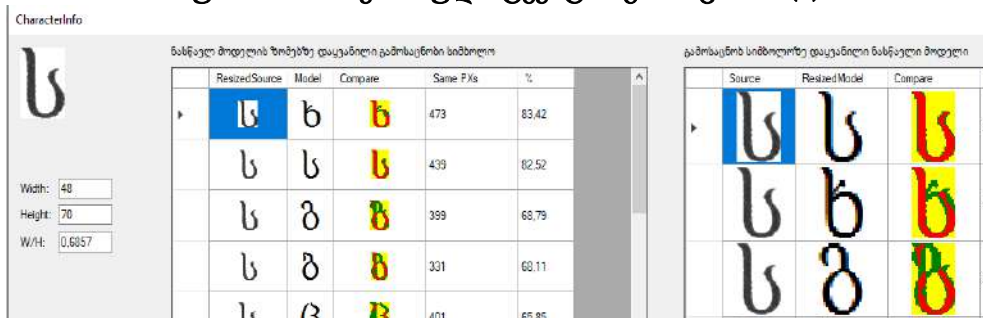
სურათი 103 - ქართული ტექსტის გამომცნობი (8)

ამის შემდეგ ალბათურ თავსებადობას ითვლის პროცენტში (პროცენტი უდრის = დამთხვეული პიქსელები * 100 / პიქსელები სულ) და ამ პროცენტული ალბათობების კლებადობით პროგრამა გვთავაზობს მონაცემებს ყველა დადარებაზე (ქართულის შემთხვევაში 33 მონაცემს).

ამგვარად, ქვემოთ, სურათი_104-დან ჩანს, რომ ამოსაცნობი სურათი არის „თ“ ასო 90.88 ალბათობით, „ო“ ასო არის 75.84 ალბათობით ... „ღ“ ასო არის 63.49 ალბათობით და ა.შ. ყველა დასწავლილი სიმბოლოთან მიმართებაში გვეძლევა პროცენტული მსგავსების მონაცემი. ასევე, სურათ_105-ზე ნაჩვენებია მარცხენა და მარჯვენა მეთოდებით მსგავსების კლებით დალაგებული სია, სადაც ორივე მეთოდი უკვე მესამე რიგიდან სხვადასხვა ალბათურ თავსებადობას იძლევიან



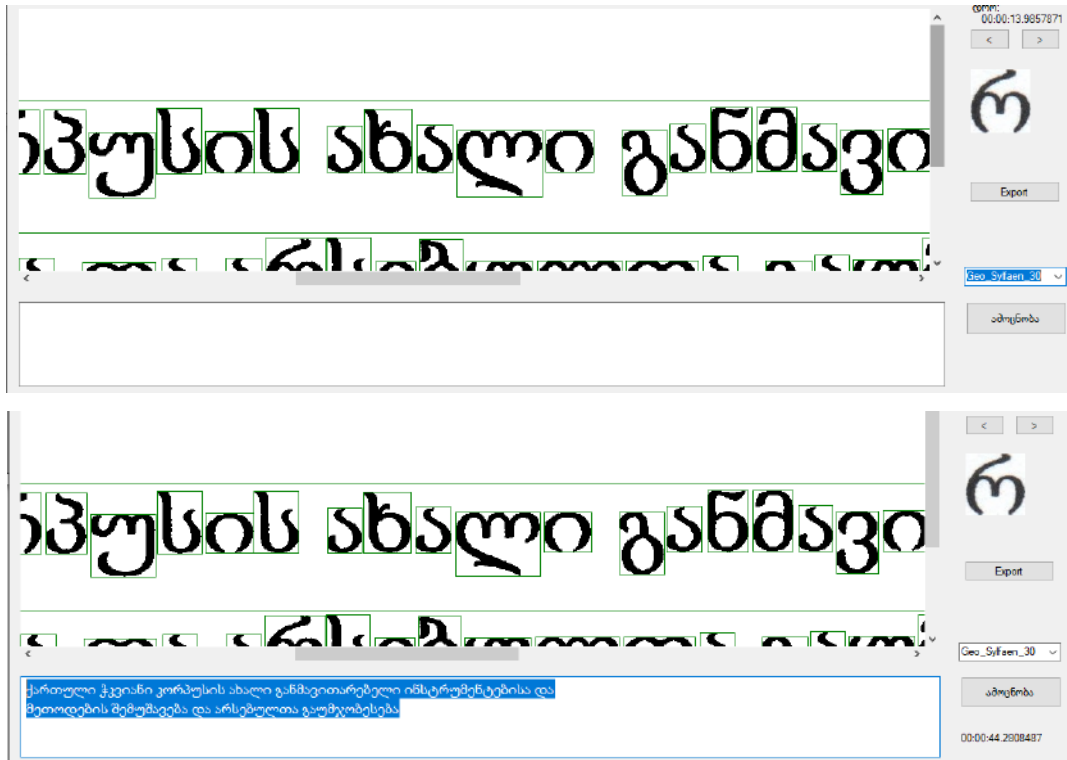
სურათი 104 - ქართული ტექსტის გამომცნობი (9)



სურათი 105 - ქართული ტექსტის ამომცნობი (10)

არის მომენტი სადაც ამ ორ მეთოდს პირველივე რიგიდან სხვადასხვა ალბათური თავსებადობა აქვთ. მაგალითად თუ ავირჩევთ ასო „ს“ სურათს ამოსაცნობად მარცხენა მოდელი გვეუბნება რომ ეს არის ასო „ბ“ ხოლო მარჯვენა მოდელი მას სწორად იცნობს (იხ. სურათი_106). რადგან უკვე გვაქვს მართკუთხედებად შემოფარგლული სურათები და ასევე დასწავლილ მოდელთან ალბათური თავსებადობით ამოცნობა, შეგვიძლია უკვე დასწავლილი სიიდან ავირჩიოთ ჩვენთვის სასურველი მოდელი, რომელსაც პროგრამა გამოიყენებს როგორც დასწავლილი მოდელი და „ამოცნობა“ ლილაკზე დაჭერით პროგრამა ყველა შემოფარგლულ სურათს გაუშვებს დასწავლილ მონაცემებთან დასადარებლად და შემდგომ პროგრამა ქვედა ტექსტის გრაფაში დაგვიწერს სურათიდან ამოცნობილ ტექსტს (იხ. სურათი_107, ზედა ნაწილი). ამოცნობილი ტექსტი არის „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების

შემუშავება და არსებულთა გაუმჯობესება“, რომელიც დალაგებულია იგივე სტრიქონული თანწყობით, რაც დასკანერებულ სურათზე იყო (იხ. სურათი_107, ქვედა ნაწილი). ჰარის საპოვნელად არის მითითება, რომ თუ ორ ასოს შორის სიგანე 10 პიქსელზე მეტია, ჩაითვალოს, რომ ამ ორ ასოს შორის ჰარია განთავსებული.



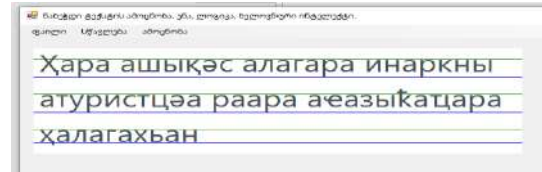
სურათი 106-107 - ქართული ტექსტის ამომცნობი (11)

2.14. აფხაზური ტექსტის ამომცნობის პირველი საცდელი სისტემები

ქვემოთ მიმოხილული აფხაზური ტექსტის ამომცნობი სისტემა აგებულია გიორგი ჩიჩუას ქართულისთვის შემუშავებულ მეთოდზე დაყრდნობით. ამ მიდგომით ტექსტის ამოსაცნობად ჯერ საჭიროა პროგრამას ვასწავლოთ ყველა ის აფხაზური სიმბოლო, რომელიც გვინდა შემდგომ ამოიცნოს, ესენია შემდეგი სიმბოლოები: **ә а к њ ҕ е е ц ц у к е н г ш з з х х ҟ ф ы в а п р л д ж ч ч с м и т ь ь ҟ**. პროგრამა ჩამოუვლის ყველა ამ სიმბოლოს და დაასურათებს რათა ქონდეს თითოეულ სიმბოლოზე გადაღებული სურათი, ამით ჩვენმა პროგრამამ უკვე ისწავლა აფხაზური შრიფტი. აფხაზური ტექსტის ამოსაცნობად მომდევნო ბიჯია რომ პროგრამას

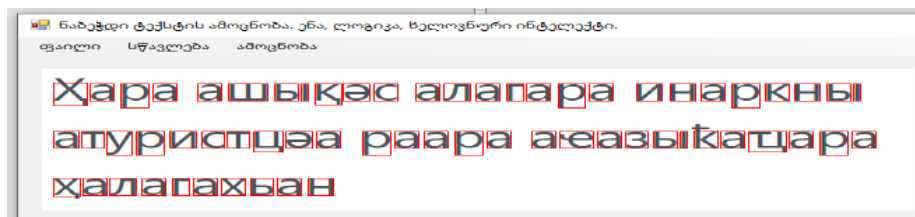
მივაწოდოთ სურათი სადაც აფხაზური ტექსტი იქნება დატანილი, მაგალითისთვის ქვემოთ სურათით მოცემული ტექსტი ავიღოთ.

Ҷара ашықәс алагара инаркны
атуристцәа раара аҕазықатцара
ҳалагахьан



სურათი 108-109 - აფხაზური ტექსტის ამომცნობი (1)

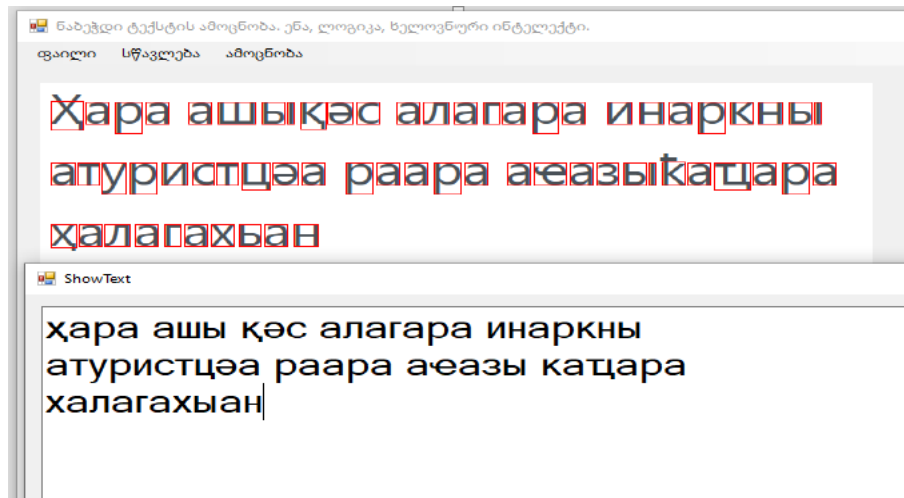
როცა მივცემთ პროგრამას ამოცნობის პროცესის დაწყებას ჯერ პროგრამა გააანალიზებს სურათს და ნახავს მასში არსებულ სტრიქონებს. შემდეგ თითოეულ სტრიქონში დაივლის და გამოარჩევს ასოებს



სურათი 110 - აფხაზური ტექსტის ამომცნობი (2)

ამის შემდეგ უკვე სურათიდან ამოიღებს თითოეულ ასოს და ზემოთ ნახსენები სწავლების რეჟიმის დროს თითოეულ ნასწავლ სიმბოლოზე გადაღებულ სურათებს დაუყვება და დაიწყებს ამოცნობის პროცესს, დადარებას ყველა ნაპოვნი სიმბოლოსთვის, სურათში ნაპოვნი სიმბოლოს სურათი რომელ ნასწავლი ასოს სურათსაც დაემთხვევა პროგრამა ჩათვლის რომ ეს სურათი სწორედ ის ასოა რომელმაც მოახდინა დადარება და ასე აკეთებს თითოეული ასოსთვის, შემდგომ ახალ ფანჯარაში გამოგვიტანს შედეგს (იხილეთ სურათი_111). პროგრამა საცდელ რეჟიმშია და მიმდინარეობს მისი ხარისხობრივი დახვეწა მაგრამ ამ კონკრეტულ შემთხვევაში ჩანს რომ ამოცნობილი ტექსტი მაღალი პროცენტულობით შეესაბამება გამოსაცნობად მიცემულ სურათზე დატანილ ტექსტს. პროგრამას ასევე გააჩნია სწავლების რეჟიმი და ჩვენ შეგვიძლია ისეთი

ამოცნობილი სიმბოლო რომელიც არა რის სწორად ამოცნობილი მივუთითოთ მისი რეალური მნიშვნელობა.



სურათი 111 - აფხაზური ტექსტის ამომცნობი (3)

ახლა აფხაზური ტექსტის კიდევ ერთი ამომცნობი სისტემა, რომელიც კონსტანტინე ფხაკაძის მეთოდზე დაყრდნობით შემუშავებული ზემოგანხილული ქართული ტექსტის ამომცნობ სისტემაზე დაყრდნობითაა აგებული. სურათ_112-ზე ნაჩვენებია აფხაზური ასოების სწავლების პროცესი, ყველა აფხაზური სიმბოლო (ესენია: а аҭ к кҭ нҭ г гҭ ҭ ц ц у к е н г ш з з х х ц ф ы в а п р о л д ж ч ч с м и т ь н т სიმბოლოები) სისტემამ ისწავლა Arial ფონტში ფონტის ზომა 30 ში და სისტემაში შენახულ იქნა სახელად „Abk_Arial_30“, ასევე ქვედა სივრცეში ხედავთ სიმბოლოების ჩამონათვალს მისი შესაბამისი დასურათების სიგანით სიმაღლით და სიგანის შეფარდება სიმაღლესთან მონაცემებით

The image shows a window titled 'სწავლება' with a 'სიმბოლოები' field containing 'а аҭ к кҭ нҭ г гҭ ҭ ц ц у к е н г ш з з х х ц ф ы в а п р о л д ж ч ч с м и т ь н т'. The 'ფონტი' is 'Arial' and 'ზომა' is '30'. Below the table is a 'გაზსნა' button. The table below shows the following data:

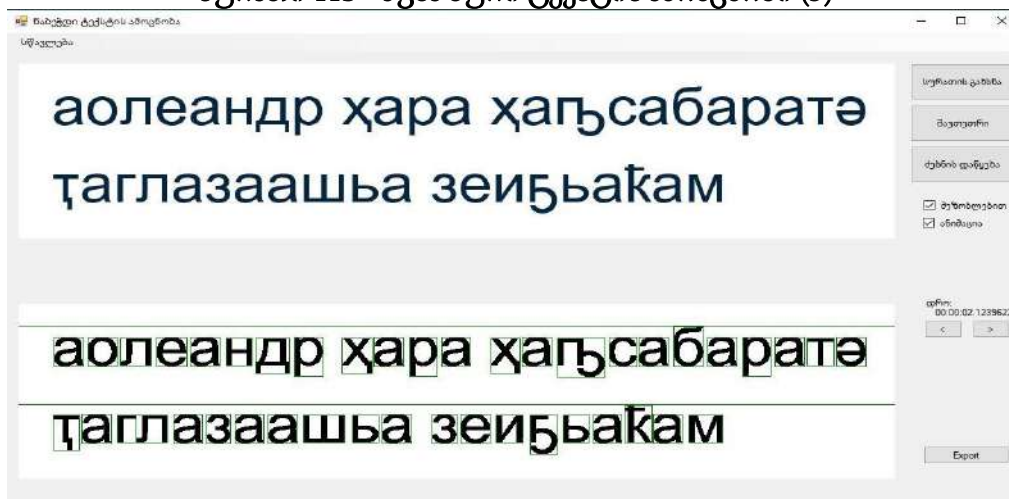
Symbol	Image	Width	Height	Width/Height
а	а	20	21	0,9524
аҭ	аҭ	20	21	0,9524
к	к	16	27	0,5926
кҭ	кҭ	18	29	0,6207
нҭ	нҭ	32	29	1,1034
г	г	16	29	0,5517

სურათი 112 - აფხაზური ტექსტის ამომცნობი (4)

ამოსაცნობი ტექსტი არის Arial ფონდით დაბეჭდილი აფხაზური ორსტრიქონიანი ტექსტი რომელიც სურათზეა

**აолеандр ҳара ҳағсабаратә
җаглазаашья зейҗьякам**

სურათი 113 - აფხაზური ტექსტის ამომცნობი (5)



სურათი 114 - აფხაზური ტექსტის ამომცნობი (6)

სისტემას მიცემული აქვს სურათ_114-ზე განთავსებული სურათი და ასოების ბარიერების ამოცნობის შემდგომ გვექნება შემდეგი მდგომარეობა: ქვედა სასურათო სივრცეში სისტემა აგვისახავს ყველა სტრიქონის დასაწყისს მწვანე ჰორიზონტალურ ზოლად და ასევე სტრიქონში შემავალ სიმბოლოებს მართკუთხედ მწვანე ფორმებში. ასევე, როგორც სურათ_114-დან ხედავთ სისტემამ ყველა სიმბოლს მონაცემები სწორად ამოიცნო და შემდგომ შემოფარგლა.ამოცნობის პროცესი ანალოგიურია როგორც ქართლის შემთხვევაში, მაგალითად სურათ_115-ზე გამოსახული ასოსთვის ამომცნობი სისტემის შედეგი გამოსახულია სურათ_116-ზე.



სურათი 115 - აფხაზური ტექსტის ამომცნობი (7)

CharacterInfo

საჩვენებელი ზომაზე დაკანონიერებული სიმბოლო

ResizedSource	Model	Compare	Same Pix	%
X	X	X	575	88.21
X	Д	X	332	69.17
X	Ц	X	222	65.87
X	У	X	387	65.62
X	Ц	X	542	64.57
X	У	X	408	62.95
X	X	X	225	62.50
X	З	X	358	60.27
X	Л	X	215	59.72

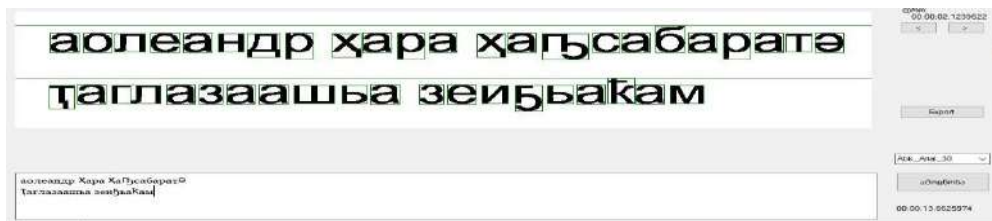
Width: 34
Height: 85
W/H: 1:2.51

აპოზიციონირებული დაკანონიერებული სიმბოლო

Source	ResizedModel	Compare	Total Pix	Same Pix	NotCompare	%
X	X	X	1564	1430	134	91
X	Д	X	1564	1089	495	68
X	Ц	X	1564	1023	541	65
X	К	X	1564	1000	564	63
X	У	X	1564	995	569	63
X	У	X	1564	995	569	63
X	Ц	X	1564	960	598	61
X	X	X	1564	929	635	59
X	К	X	1564	926	638	59

სურათი 116 - აგზაზური ტექსტის ამომცნობი (8)

სისტემა ამომცნობის შემდგომ დაგვიბრუნა ტექსტი (ეს ტექსტი არის - аолеандр хара хатсабаратә таглазаашья зейџакам (იხ., სურათი_117)), რაც სრულ თანმთხვევაშია ამოსაცნობ ტექსტთან



სურათი 117 - აგზაზური ტექსტის ამომცნობი (9)

2.15. ქართული სიტყვების სრული ავტომატური ანალიზატორის (ანუ ტაგერის, დესკრიპტორის, ლემატიზატორის, გენერატორისა და ანოტატორის) პირველი საცდელი ვერსია

ჯერ განვიხილოთ ასეთი სისტემის ასაგებად შემუშავებული ახალი გაძლიერებული ინსტრუმენტები.

კვანტორი

სახელი

სახელი

სახელი

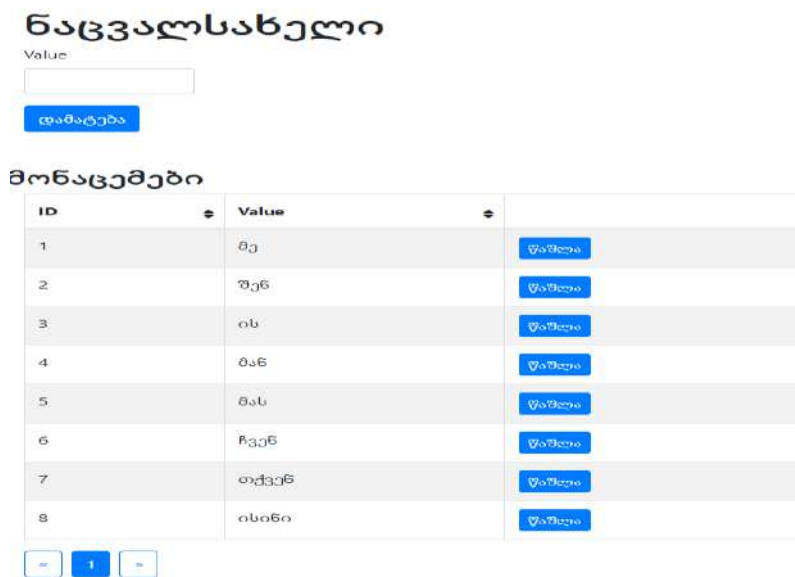
დასრულა

მონაცემები

ID	Value	Value0	Value1	Value2
1	ვფაშა	ვფაშა	ვფაშა	დასრულა
2	ფოჯა	ფოჯა	ფოჯა	დასრულა
3	ფოჯგო	ფოჯგო	ფოჯგო	დასრულა
4	ფოჯგო	ფოჯგო	ფოჯგო	დასრულა

სურათი 118 - კვანტორის ამომცნობი

კორპუსი გადლიერებულია ისეთი ტიპის სიტყვებით როგორც არის კვანტორი. კორპუსში ჩაყრილია ყველა კვანტორის ფორმა, რისი მეშვეობითაც უკვე შეუძლია კორპუსს დაადგინოს შემოსული სიტყვის ტიპი არის თუ არა კვანტორი. ზემოთ სურათზე ჩანს თუ როგორ მოხდა კვანტორების დამატება მარტივად , შეივსო კვანტორის სამი ფორმა და ღილაკზე „დამატება“ დაჭერით მითითებული ფორმები ჩაემატა კორპუსში ქვემოთ მოცემულია უკვე არსებული კვანტორების სია.



სურათი 119 - ნაცვალსახელის ამომცნობი

კორპუსში ასევე არის ნაცვალსახელის ტიპის სიტყვების შესაყვანი პანელი, სადაც შეგვიძლია ნაცვალსახელი ჩავწეროთ და ღილაკზე ერთი დაჭერით კორპუსმა ისწავლოს ეს ნაცვალსახელი და უკვე შეეძლება დანერგილ ინსტრუმენტებში გამოიყენოს სიტყვის ფორმა არის თუ არა ნაცვალსახელი.

სურათ_120-ზე ხედავთ სისტემას სადაც შეიძლება ნაცვალსახელის დამატება. ქვემოთ უკვე დამატებული ნაცვალსახელების სია ჩანს სადაც შეგვიძლია წავშალოთ ნაცვალსახელი თუ არასწორი ფორმით იქნა დამატებული.

კორპუსის მონაცემებს ასევე აქვს საშუალება რომ მივუთითოთ კავშირების სია, შესაძლებელია კავშირების დამატება და არსებულის წაშლა.

კავშირი

Value

დამატება

მონაცემები

ID	Value	
3	და	წაშლა
4	თუ	წაშლა
5	მაგრამ	წაშლა
6	ხოლო	წაშლა
7	თურმე	წაშლა
8	ან	წაშლა
9	რომ	წაშლა
10	რატა	წაშლა

« 1 »

სურათი 120 - კავშირების ამომცნობი

სურათ_120-ზე ჩანს კავშირის დამატებისთვის საჭირო ფორმა და ღილაკზე „დამატება“ დაჭერის ეს ფორმა ჩავარდება მონაცემების სიაში. ქვემოთ ვხედავთ კავშირების იმ ჩამონათვალს რომლებიც უკვე არის დამატებული კორპუსში. კორპუსს ასევე შეგვიძლია ვასწავლოთ ისეთი სიტყვის ფორმები რომლებიც არიან შორისდებულები

შორისდებული

Value

დამატება

მონაცემები

ID	Value	
1	ო	წაშლა
2	უეუპ	წაშლა
3	კაი	წაშლა
4	ან	წაშლა
5	ონ	წაშლა
6	პო	წაშლა
7	კა	წაშლა
8	კიშ	წაშლა
9	აოჩაო	წაშლა
10	იპ	წაშლა

« 1 »

რიცხვითი სახელი

Value

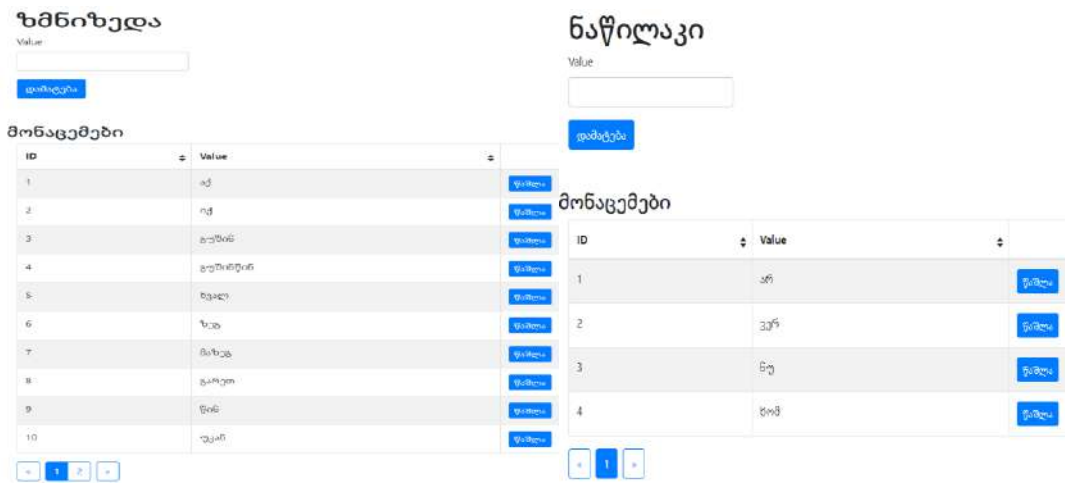
დამატება

მონაცემები

ID	Value	
1	კრთი	წაშლა
2	ოთი	წაშლა
3	სამი	წაშლა
4	ოთხი	წაშლა
5	ხუთი	წაშლა
6	ექვთი	წაშლა
7	შვიდი	წაშლა
8	ჩხა	წაშლა
9	ცხრა	წაშლა
10	ათი	წაშლა

სურათი 121-122 - რიცხვითი სახელისა და შორისდებულის ამომცნობი

სურათ_121-ზე ასახულია სისტემა სადაც შეგვიძლია შევიტანოთ შორისდებულის ისეთი ფორმა რომელიც ჯერ კორპუსში არარის ჩამატებული. „დამატება“ ღილაკზე დაჭერით ჩავამატოთ შორისდებულების მონაცემებში უკვე ახალი შორისდებული რომელიც გამოჩნდება სურათ_121-ზე ქვევით ნაჩვენებ შორისდებულების სიაში სადაც შეგვიძლია დავათვალიეროთ ყველა ის შორისდებული, რომელიც უკვე არის დამატებული მონაცემთა ბაზაში ან სურვილისამებრ წავშალოთ არასწორი ჩანაწერი. რიცხვითი სახელის სწავლის პროცესი კორპუსში ჯერ-ჯერობით მარტივია, ვწერთ რიცხვით სახელს და ვამატებთ სიაში (სურათი_122).



სურათი 123-124 - ზმნიზედასა და ნაწილაკის ამომცნობი

ზმნიზედების (სურათი_123) და ნაწილაკების (სურათი_124) სწავლების ზემოგანხილული აღნიშნული სახელების სწავლების პროცესის. ანალოგიურია. ანუ, აქაც, ვუთითებთ სასურველი ზმნიზედასა თუ ნაწილაკის ფორმას და დამატება ღილაკზე გააქტიურებით ვაფართოებთ კორპუსში არსებულ ზმნიზედებისა და ნაწილაკების სიას.

ჯერ-ჯერობით ანალოგიურად ვიქცევით ზმნის სწავლების პროცესისათვისაც (სურათი_125), თუმცა, ზმნის სწავლება აგრეთვე ხდება ავტომატურად გენერატორების მეშვეობით.

ზმნა

Value

დამატება

მონაცემები

ID	Value	
1	ვსწავლობ	წაშლა
2	ვაკეთებ	წაშლა
3	ხატავს	წაშლა
4	შიდისა	წაშლა
5	შიფრინავს	წაშლა
6	დღას	წაშლა

← + →

სურათი 125 - ზმნის ამომცნობი

ახლა უკვე ძალიან მოკლედ თავად ქართული სიტყვების სრული ავტომატური ანალიზატორის პირველი საცდელ ვერსიზე.¹⁷

ჩვენ მიერ უკვე შემუშავებული ქართული სიტყვების ანალიზატორი არის სისტემა რომელიც შეტანილ სიტყვაზე ადგენს არის თუ არა კვანტორი, კავშირი, ნაცვალსახელი, შორისდებული, არსებითი სახელი, ზედსართავი სახელი, ზმნა თუ არცერთი ჩამოთვლილთაგან.

არცერთი 🔍

არცერთი	
ფიქსი 1	არცერთი
ფიქსი 2	არცერთი
კუჭმუცაღი	კი
ტაბი	კვანტორი
ფიქსის ტაბი	ი

თქვენ 🔍

თქვენ	
ფიქსი 1	თქვენ
ფიქსი 2	თქვენ
კუჭმუცაღი	არა
ტაბი	ნაცვალსახელი
ფიქსის ტაბი	ი

სურათი 126-127 - კვანტორის და ნაცვალსახელის ამომცნობი

¹⁷ ეს უნიკალური სისტემა არის ერთგვარი კომბინირებული მოცემულობა ქართული სიტყვების სრული ტაგერის, დესკრიპტორის, ლემატიზატორის, გენერატორისა და ანოტატორის.

ბიჭოს	
ფურცე 1	ბიჭო
ფურცე 2	ბიჭო
კუჭმუცა	არა
ტიპი	მორისდებულის
ფურცის ტიპი	ო

მაგრამ	
ფურცე 1	მაგარა
ფურცე 2	მაგრ
კუჭმუცა	კი
ტიპი	კავშირი
ფურცის ტიპი	ა

სურათი 128-129 - მორისდებულის და კავშირის ამომცნობი

მწვანე	
ფურცე 1	მწვანე
ფურცე 2	მწვანე
კუჭმუცა	არა
ტიპი	ზედსართავი სახელი
ფურცის ტიპი	შ
სულ	29612

მცენარე	
ფურცე 1	მცენარე
ფურცე 2	მცენარე
კუჭმუცა	არა
ტიპი	არსებითი სახელი
ფურცის ტიპი	შ

სურათი 130-131 - ზედსართავის და არსებითის ამომცნობი

შესაძლო მრავალტიპიანი სიტყვის ამოცნობა მაგალითად სიტყვა „და“ს შემთხვევაში სისტემა გვეუბნება რომ ის არის კავშირიც და ამავდროულად არსებითი სახელიც (იხილეთ სურათი_132 ქვემოთ).

და	
ფურცე 1	და
ფურცე 2	და
კუჭმუცა	არა
ტიპი	კავშირი, არსებითი სახელი
ფურცის ტიპი	ა

მესამე	
ფურცე 1	მესამე
ფურცე 2	მესამე
კუჭმუცა	კი
ტიპი	რიცხვითი სახელი
ფურცის ტიპი	ე

სურათი 132-133 - მრავალტიპიანი სიტყვების ამომცნობი

ფეხნი		ვერ	
ფეხი 1	გეშნი	ვერ	
ფეხი 2	გეშნი	ვერ	
კენშეაფი	არა	არა	
ტიპი	შმნიღლა	ნანიღლაკი	
ფეხის ტიპი	"	o	
სულ			1755

სურათი 134-135 - ზმნის და ნაწილაკის ამომცნობი

ვაკეაუბ	
ფეხი 1	ვაკეა
ფეხი 2	ვაკეა
კენშეაფი	არა
ტიპი	შმნი
ფეხის ტიპი	ა

სურათი 136 - ზმნის ამომცნობი

ასევე, ავტომატური მეთოდით ზმნის ამომცნობისთვის საჭიროა საძიებო ველში შევიტანოთ ზმნის ფორმა და დავაჭიროთ „ანალიზი“ ღილაკს.

სისტემა პოულობს შეტანილი სიტყვისთვის სავარაუდო ფუმეს და შემდგომ სვამს შემდეგ ოთხ ბლოკში: 1; მხოლობითი-მხოლობითი; 2. მხოლობითი-მრავლობითი; 3. მრავლობითი-მხოლობითი; 4. მრავლობითი-მრავლობითი. ამის შემდეგ სისტემა აგენერირებს ყველა ამ დროისათვის ფორმებს, რაც სურათ_137-ზეა ნაჩვენები, შემდგომ იღებს ამ ფორმების სინტაქსურ მახასიათებლებს. ქვემოთ სურათ_137-ზე მოყვანილია “ვწერ” ზმნის მაგალითი სადაც გამოიცნო სიტყვა, როგორც ზმნა და, ასევე, მიხვდა იმას, რომ ის არის ორ უღლებად ადგილიანი. გარდა ამისა, სისტემა მიხვდა იმასაც, რომ მას ორი პრინციპულად განსხვავებული სინტაქსური ფორმულა აქვს. ესენია: 1. ვწერ [მე] [მას/ნს] [ნს/ნებს]; 2. ვწერ [მე] [მათ/ნს] [ნს/ნებს].

ვწერ detect ანალიზი

1. ფუძე: წერ

ვწერ მე [მას/ნს] [ნს/ნებს]

ვწერ მე [მათ/ნებს] [ნს/ნებს]

მხოლობითი-მხოლობითი

#	წესი	ფორმა	რაოდენობა	სიხშირე
1.	ვ[R1]	ვწერ	4491	0.0009332655879355006
2.	გ[R1]	გწერ	312	0.00006483608626940017
3.	[R1]	წერ	793	0.00016479171926805878
4.	მ[R1]	მწერ	152	0.000031586811259451366
5.	[R1]ს	წერს	91813	0.019079473040552686
6.	გ[R1]ს	გწერს	142	0.000029508731571329566
7.	მ[R1]ს	მწერს	783	0.00016271363957993698

მხოლობითი-მრავლობითი

#	წესი	ფორმა	რაოდენობა	სიხშირე
1.	ვ[R1]	ვწერ	4491	0.0009332655879355006
2.	გ[R1]თ	გწერთ	428	0.00008894181065161306
3.	[R1]	წერ	793	0.00016479171926805878
4.	გვ[R1]	გვწერ	1	2.0780796881218005e-7
5.	[R1]ს	წერს	91813	0.019079473040552686
6.	გ[R1]თ	გწერთ	428	0.00008894181065161306
7.	გვ[R1]ს	გვწერს	116	0.000024105724382212885

ორ უღლებად ადგილიანი

მრავლობითი-მხოლობითი

#	წესი	ფორმა	რაოდენობა	სიხშირე
1.	ვ[R1]თ	ვწერთ	531	0.0001103460314392676
2.	გ[R1]თ	გწერთ	428	0.00008894181065161306
3.	[R1]თ	წერთ	551	0.0001145021908155112
4.	მ[R1]თ	მწერთ	67	0.000013923133910416063
5.	[R1]ენ	წერენ	5709	0.0011863756939487359
6.	გ[R1]ენ	გწერენ	55	0.000011429438284669903
7.	მ[R1]ენ	მწერენ	237	0.00004925048860848667

მრავლობითი-მრავლობითი

#	წესი	ფორმა	რაოდენობა	სიხშირე
1.	ვ[R1]თ	ვწერთ	531	0.0001103460314392676
2.	გ[R1]თ	გწერთ	428	0.00008894181065161306
3.	[R1]თ	წერთ	551	0.0001145021908155112
4.	გვ[R1]თ	გვწერთ	16	0.000003324927500994881
5.	[R1]ენ	წერენ	5709	0.0011863756939487359
6.	გ[R1]ენ	გწერენ	55	0.000011429438284669903
7.	გვ[R1]ენ	გვწერენ	96	0.000019949565005969282

სურათი 137 - ზმნის ამომცნობი (1)

ქვედა სურათ_138-ზე მითითებულია სიტყვა “ვცხოვრობ”, რომელზეც სისტემამ დაასკვნა, რომ ის არის ერთ უღლებად ადგილიანი ზმნა [13].

ვცხოვრობ detect ანალიზი

1. ფუძე: ცხოვრობ

ვცხოვრობ მე

მხოლობითი-მხოლობითი

#	წესი	ფორმა	რაოდენობა	სიხშირე
1.	ვ[R1]	ვცხოვრობ	5149	0.001070003231413915
2.	გ[R1]	გცხოვრობ	0	0
3.	[R1]	ცხოვრობ	2043	0.0004245516802832838
4.	მ[R1]	მცხოვრობ	13	0.0000027015035945583404
5.	[R1]ს	ცხოვრობს	33641	0.006990867878810549
6.	გ[R1]ს	გცხოვრობს	0	0
7.	მ[R1]ს	მცხოვრობს	0	0

მხოლობითი-მრავლობითი

#	წესი	ფორმა	რაოდენობა	სიხშირე
1.	ვ[R1]	ვცხოვრობ	5149	0.001070003231413915
2.	გ[R1]თ	გცხოვრობთ	0	0
3.	[R1]	ცხოვრობ	2043	0.0004245516802832838
4.	გვ[R1]	გვცხოვრობ	0	0
5.	[R1]ს	ცხოვრობს	33641	0.006990867878810549
6.	გ[R1]თ	გცხოვრობთ	0	0
7.	გვ[R1]ს	გვცხოვრობს	0	0

ერთ უღლებად ადგილიანი

მრავლობითი-მხოლობითი

#	წესი	ფორმა	რაოდენობა	სიხშირე
1.	ვ[R1]თ	ვცხოვრობთ	5766	0.0011982207481710302
2.	გ[R1]თ	გცხოვრობთ	0	0
3.	[R1]თ	ცხოვრობთ	1046	0.00021736713537754034
4.	მ[R1]თ	მცხოვრობთ	0	0
5.	[R1]ენ	ცხოვრობენ	12631	0.0026248224540666646
6.	გ[R1]ენ	გცხოვრობენ	0	0
7.	მ[R1]ენ	მცხოვრობენ	0	0

მრავლობითი-მრავლობითი

#	წესი	ფორმა	რაოდენობა	სიხშირე
1.	ვ[R1]თ	ვცხოვრობთ	5766	0.0011982207481710302
2.	გ[R1]თ	გცხოვრობთ	0	0
3.	[R1]თ	ცხოვრობთ	1046	0.00021736713537754034
4.	გვ[R1]თ	გვცხოვრობთ	0	0
5.	[R1]ენ	ცხოვრობენ	12631	0.0026248224540666646
6.	გ[R1]ენ	გცხოვრობენ	0	0
7.	გვ[R1]ენ	გვცხოვრობენ	0	0

სურათი 138 - ზმნის ამომცნობი (2)

სიტყვა, რომელიც სისტემამ არ იცის ან ჯერ არ აქვს ნასწავლი, მასზე წერს რომ იგი არის ამოუცნობი როგორც ნაჩვენებია ქვედა სურათზე [13].

მეცხრაშეტე	
ფუძე 1	მეცხრაშეტე
ფუძე 2	მეცხრაშეტე
კუშმვალი	კი
ტიპი	ამოუცნობი
ფუძის ტიპი	ე

სურათი 139 - ამოუცნობი სიტყვის ამომცნობი

2.16. სადისერტაციო კვლევის მთავარი შედეგი ანუ ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია

ამგვარად, ახლა უკვე, ძალიან მოკლედ სადისერტაციო კვლევის მთავარ შედეგზე ანუ მოსაუბრე ქართული ინტელექტუალური ანუ ჭკვიანი ვებ-კორპუსის გაძლიერებისა და გაფართოების გზით შემუშავებულ ქართულ ანუ იბერიულ-კავკასიურ უნივერსალურ ჭკვიან კორპუსზე (<https://corpus.ge/>), რომელშიც ჩადგმული და/ან რომლის განვითარების მიზნით შემუშავებული უნიკალური ქართული და აფხაზური ენობრივი ტექნოლოგიური სისტემები ჩვენ ზემოთ უკვე მიმოვიხილეთ. ამასთან, მიუხედავად იმისა, რომ ქართული უნივერსალური ჭკვიანი ვებ-კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია ჯერ მხოლოდ ერთი ნაბიჯია იბერიულ-კავკასიური უნივერსალური ჭკვიანი კორპუსის სრული სახით აგებისაკენ, ის საცდელი სახით უკვე:

1. არის ერთადერთი იბერიულ-კავკასიური თვითგანვითარებადი კორპუსი - ის თავადვე ანვითარებს თავის მოცულობით და ინტელექტუალურ მაჩვენებლებს.¹⁸

¹⁸ კორპუსში ყოველწუთიერ ავტომატურად შემოდის ახალ-ახალი იბერიულ-კავკასიური ტექსტობრივი მონაცემები. ამასთან, კორპუსში ჩადგმული ინტელექტუალური

2. არის ერთადერთი ინტერაქტიური კორპუსი ქართულ და აფხაზურ სასაუბრო ენაში. - მასში უკვე საცდელი სახით ჩადგმულია ქართული და აფხაზური ხმოვანი მართვის მოდული.

3. არის ერთადერთი იბერიულ-კავკასიური მრავალენოვანი კორპუსი - მასში უკვე ჩადგმული სახით ფუნქციონირებს თვითგანვითარებადი ქართული, აფხაზური, ჩეჩნური, ყაბარდოული, ლეკური და მეგრული კორპუსები.

4. არის ერთადერთი მრავალმოდულური - მასში უკვე ჩადგმულია საცდელი სახით მომუშავე ქართული ტიტრირებული მეტყველების კორპუსი და ამავე ტიპის აფხაზური, ჩეჩნური, ყაბარდოული, ლეკური და მეგრული კორპუსების ამგები და დამამუშაველი ინსტრუმენტები, რაც ჩვენ გვჭირდება ქართული, აფხაზური და სხვა იბერიულ-კავკასიური სამეტყველო ენების მასინთეზებელი და ამომცნობი სისტემების აგებისა და შემდგომი ხარისხობრივი გაუმჯობესებისათვის.

5. არის ყველა მოცულობითი ქართული კორპუსი. - დღევანდელი ანუ 2019 წლის 18 დეკემბრის მონაცემებით კორპუსი მთლიანობაში მოიცავს 331782593 სიტყვა-ტოკენს, მათ შორის 6104732 განხვავებულია, რაც იმას ნიშნავს, რომ ორივე ამ მაჩვენებლით ჩვენი კორპუსი ყველაზე მოცულობითი ქართული კორპუსია.

იმაში, რომ ეს მართლაც ასეა, გვარწმუნებს <http://gnc.gov.ge/gnc/page> მისამართზე განთავსებული ვებ-გვერდი „ქართული ენის ეროვნული კორპუსი“ და <http://gnc.gov.ge/gnc/corpus-list?session-id=248210997162307> მისამართზე განთავსებული მასალა, სადაც თავმოყრილია დღეს მოქმედი/არსებული ქართული კორპუსების მეტი წილი.¹⁹ ესენია:

პროცედურები პერმანენტულად აფართოებს კორპუსის ცოდნას ამ მასში უკვე შემოსული ტექსტებით მოცემულ შინაარსებზე.

¹⁹ ქართულ ვიკიპედიაში იხილეთ სტატია „ქართული ენის ეროვნული კორპუსი“, რომლის თანახმადაც ამ პროექტის ფარგლებში „1999 წლიდან 2012 წლამდე 8 ჯგუფური საერთაშორისო სამეცნიერო პროექტი და 7 ქართველი მეცნიერის ინდივიდუალური

1. GNC ანუ ქართული ენის ეროვნული კორპუსი (ძველი ქართული), რომლის ჯამური ტოკენური მოცულობაა 6062122 (2019.12.18);
2. GNC ანუ ქართული ენის ეროვნული კორპუსი (საშუალო ქართული), რომლის ჯამური ტოკენური მოცულობაა 1432262 (2019.12.18);
3. GNC ანუ ქართული ენის ეროვნული კორპუსი (ახალი ქართული), რომლის ჯამური ტოკენური მოცულობაა 2108370 (2019.12.18);
4. GRC ქართული რეფერენციალური კორპუსი, რომლის ჯამური ტოკენური მოცულობაა 202728329 (2019.12.18);
5. GDC ანუ ქართული დიალექტური კორპუსი, რომლის ჯამური ტოკენური მოცულობაა 1694362 (2019.12.18);
6. SSGG ანუ სოციოლინგვისტური სიტუაცია თანამედროვე საქართველოში, რომლის ჯამური ტოკენური მოცულობაა 152708 (2019.12.18);
7. GNC ანუ პოლიტიკური ტექსტები, რომლის ჯამური ტოკენური მოცულობაა 1436075 (2019.12.18);
8. GNC იურიდიული ტექსტები, რომლის ჯამური ტოკენური მოცულობაა 1436075 (2019.12.18);
9. GNC ანუ ქართული ენის ეროვნული კორპუსი (მეგრული), რომლის ჯამური ტოკენური მოცულობაა 89404 (2019.12.18);
10. GNC ანუ ქართული ენის ეროვნული კორპუსი (სვანური), რომლის ჯამური ტოკენური მოცულობაა 473180 (2019.12.18);

ანუ, 2019 წლის 18 დეკემბრის მონაცემებით ჩვენი კორპუსი ანუ ქართული უნივერსალური ჭკვიანი კორპუსი, რომელიც ქართული ენის ეროვნული კორპუსისაგან განსხვავებით მხოლოდ ოთხი წელია შენდება, ეროვნულ კორპუსში შემავალი კორპუსების ჯამურ მოცულობას, რაც

პროექტი განხორციელდა იოსტ გიპერტისა და მანანა თანდაშვილის ხელმძღვანელობით, რომლის საერთო ღირებულება 1, 5 მილიონ ევროს შეადგენს. “ არადა წლებში ჩვენ კორპუსზე 2015 – 2019 გაღებული სულ რაღაც 270000 ლარი, რაც, ცხადია, ზღაპრულ წვეთია 1.5 მილიონ ევროსთან შედარებით.

217672797 სიტყვა-ტოკენია, აღემატება 331782593 - 217672797 = 114,109,796 სიტყვა-ტოკენით [13].

უფრო სრული სურათისათვის მოგვყავს ილიას სახელმწიფო უნივერსიტეტის ქართული ენის კორპუსი, რომელიც <http://corpora.iliauni.edu.ge/?q=ka/search-words> მისამართზეა განთავსებული. ამ აქ მინიშნებულ მისამართებზე განთავსებული ინფორმაციის თანახმად, ეს კორპუსი, რომელიც 2009-2016 წლებში იქმნებოდა, ჯამურად შეიცავს 100 მილიონ სიტყვა ტოკენს. ეს იმას ნიშნავს, რომ ჩვენი კორპუსი, რომელიც, როგორც უკვე აღვნიშნეთ, ჯერ მხოლოდ 4 წელია უკვე შეიცავს 331782593 სიტყვას და, შესაბამისად, ილიას სახელმწიფო უნივერსიტეტის კორპუსს მოცულობით 3,33-ჯერ ანუ, ამ შემთხვევაში, 231782593 სიტყვა-ტოკენით აღემატება [13].

ამგვარად, ჩვენი თვითგანვითარებადი ანუ მოცულობის მზრიაც ყოველწუთიერ ავტომატურად მზარდი ქართული უნივერსალური ჭკვიანი კორპუსი, რომლის აგებაზე ბევრად უფრო ნაკლები დრო და რესურსია დახარჯული ვიდრე ყველა სხვა ზემოთ მიმოხილულ ქართულ კორპუსზე და, ამავდროულად, რომლის მთავარი უპირატესობა ნამდვილად არაა მასში ჯამურად თავმოყრილი მასალის მოცულობითი მახასიათებელი, ამ მახასიათებლითაც კი საგრძნობლივ აჭარბებს სხვა დღეს არსებული ქართული კორპუსებიდან ნებისმიერს. ეს კი, თავის მხრივ, გასაგებს ხდის, რომ საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის კორპუსი ანუ ქართული უნივერსალური ჭკვიანი ვებ-კორპუსი და მასში ჩადგმული და მისი განვითარების მიზნით შემუშავებული უნიკალური ქართული ტექნოლოგიური სისტემები, რომელთაგან ნაწილი ზემოთ მიმოვიხილეთ, ციფრული კვდომის იმ მაღალი ხარისხის საფრთხეების გათვალისწინებით, რომელთა წინაშეცაა სხვა ამ მხრივ დაწინაურებულ ენებთან შედარებით მათემატიკურად არასაკმარისად შესწავლილი და, შესაბამისად, არასაკმარისად ტექნოლოგიზებული ქართული და მით უფრო აფხაზური


ენა, უკიდურესად მნიშვნელოვანია. გარდა ამისა, ეს ანუ ჩვენს მიერ უკვე შემუშავებული ქართული ანუ იბერიულ-კავკასიური უნივერსალური ჭკვიანი კორპუსი გარდა ქართული და აფხაზური ენებისა მეტად მნიშვნელოვანია აგრეთვე არამხოლოდ კორპუსში არსებული არამედ ყველა სხვა იბერიულ-კავკასიური ენის ციფრული კვდომის საფრთხისგან დაცვის თვალსაზრისით [13].

დასასრულს, წარმოვადგენთ ქართულ უნივერსალურ ჭკვიან კორპუსში ფუნქციონირებადი სრული, და ქართველური და კავკასიური ენების კორპუსებს მათი მოკლე სტატისტიკური აღწერებითურთ (იხილეთ, ქვემოთ, სურათი_140, სურათი_141, სურათი_142, სურათი_143, სურათი_144, სურათი_145, სურათი_146).


სურათი 140 - სრული კორპუსის მონაცემები

 <p>იბერიულ-კავკასიური ენები: აფხაზური ჩეჩნური ყაბარდოული ლეკური მეგრული</p>	<p>სულ შემოსული: 331,782,593 განსხვავებული: 6,104,732 ბოლო დღეს: 5,406,402</p>
<input type="text" value="შეიყვანეთ ფრაზა"/> <input type="submit" value="🔍"/>	

სურათი 141 - ქართული კორპუსის მონაცემები

 <p>იბერიულ-კავკასიური ენები: აფხაზური ჩეჩნური ყაბარდოული ლეკური მეგრული</p>	<p>სულ შემოსული: 321,303,632 განსხვავებული: 4,817,761 ბოლო დღეს: 11,089</p>
<input type="text" value="შეიყვანეთ ფრაზა"/> <input type="submit" value="🔍"/>	


სურათი 142 - აფხაზური კორპუსის მონაცემები

 <p>იბერიულ-კავკასიური ენები: აფხაზური ჩეჩნური ყაბარდოული ლეკური მეგრული</p>	<p>სულ შემოსული: 7,650,858 განსხვავებული: 817,872 ბოლო დღეს: 5,395,313</p>
<input type="text" value="შეიყვანეთ ფრაზა"/> <input type="submit" value="🔍"/>	


სურათი 143 - ჩეჩნური კორპუსის მონაცემები

 <p>იბერიულ-კავკასიური ენები: აფხაზური ჩეჩნური ყაბარდოული ლეკური მეგრული</p>	<p>სულ შემოსული: 122,540 განსხვავებული: 17,905 ბოლო დღეს: 0</p>
<input type="text" value="შეიყვანეთ ფრაზა"/> <input type="submit" value="🔍"/>	


სურათი 144 - ყაბარდოული კორპუსის მონაცემები

 <p>იბერიულ-კავკასიური ენები: აფხაზური ჩეჩნური ყაბარდოული ლეკური მეგრული</p>	<p>სულ შემოსული: 126,888 განსხვავებული: 65,532 ბოლო დღეს: 0</p>
<input type="text" value="შეიყვანეთ ფრაზა"/> <input type="submit" value="🔍"/>	

სურათი 145 - ლეკური კორპუსის მონაცემები

 <p>იბერიულ-კავკასიური ენები: აფხაზური ჩეჩნური ყაბარდოული ლეკური მეგრული</p>	<p>სულ შემოსული: 402,113 განსხვავებული: 130,483 ბოლო დღეს: 0</p>
<input type="text" value="შეიყვანეთ ფრაზა"/> <input type="submit" value="🔍"/>	

სურათი 146 - მეგრული კორპუსის მონაცემები

 <p>იბერიულ-კავკასიური ენები: აფხაზური ჩეჩნური ყაბარდოული ლეკური მეგრული</p>	<p>სულ შემოსული: 967,521 განსხვავებული: 182,840 ბოლო დღეს: 0</p>
<input type="text" value="შეიყვანეთ ფრაზა"/> <input type="submit" value="🔍"/>	

დასკვნა

მოკლე დასკვნითი შეჯამება

ქართული ენის, მით უფრო აფხაზური ენის ენობრივი რესურსებითა და ტექნოლოგიებით მხარდაჭერის მოცულობა და ხარისხი საგანგაშოდ ჩამორჩება იმ ამ მხრივ არასაკმარისად მხარდაჭერილი ევროპული ენებიდან თითქმის ნებისმიერის მოცულობასა და ხარისხს, რომლებიც ევროსტრუქტურების დაფინანსებით მეტა-ქსელის მიერ 2010 – 2012 წლებში განხორციელებული ორწლიანი საერთო ევროპული კვლევის „ევროპის ენების ციფრულ ეპოქაში“ თანახმად ციფრული კვდომის საფრთხის ქვეშ არიან დღეს უკვე კარს მომდგარ და სანახევროდ უკვე ამოქმედებული ციფრულ ეპოქაში. აღნიშნულის გათვალისწინებით, წინამდებარე სადოქტორო კვლევა „ქართული ჰკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“, რომელიც არის საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის გრძელვადიანი პროექტების „ქართული ენის ტექნოლოგიური ანბანი“ და „აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის გეგმა-პროგრამა“ კიდევ ერთი მნიშვნელოვანი ქვეპროექტი, მიმართულია საქართველოს სახელმწიფო ენების - ქართულისა და აფხაზურის სრული ტექნოლოგიური უზრუნველყოფის მიზნების მიღწევაზე. ეს სხვა სიტყვებით, ნიშნავს იმას, რომ სადოქტორო კვლევა მიმართულია ქართული და აფხაზური ენების ციფრული კვდომის საფრთხისგან დაცვის მიზნების გათვალისწინებით ქართული და აფხაზური ენების ტექნოლოგიური ანბანების აგების ამოცანების სრულ და ამომწურავ გადაწყვეტებზე, რაც, მთლიანობაში, მიმართულია ერთიანი სახელმწიფოს ფარგლებში ქართველებისა და აფხაზების ევროკავშირში, უფრო ზოგადად კი, მომავლის კულტურულ სამყაროში ქართული და აფხაზური ენებით ღირსეული ანუ სრულფასოვანი შესვლის ეროვნულ-კულტურული მნიშვნელობების მქონე მიზნების უზრუნველყოფაზე [12].

მართლაც, სადისერტაციო კვლევით კ.ფხაკაძის ქართული ენის ლოგიკურ გრამატიკაზე დაყრდნობით წინა წლებში უკვე შემუშავებული ფხაკაძე-ჩიქვინიძისა და ფხაკაძე-ჩიჩუას მეთოდებზე და, ასევე, სხვადასხვა დღეს კარგად ცნობილი ინსტრუმენტებსა და პლატფორმებზე დაყრდნობით, ერთი მხრივ, გაძლიერდა, მეორე მხრივ კი, შეიქმნა ქართული და აფხაზური ენების სრული ტექნოლოგიური უზრუნველყოფისთვის ანუ ქართული აფხაზური ენების სრულყოფილად მცოდნე ინტელექტუალური კომპიუტერული სისტემების შემუშავებისთვის, რაც იგივეა, ქართული და აფხაზური ენების ტექნოლოგიური ანბანის აგებისთვის აუცილებელი მთელი რიგი დამხმარე ინსტრუმენტები და ასეთი ინსტრუმენტების ამგები მეთოდები და ამ გზით შემუშავდა ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია, რაც სადისერტაციო კვლევის მთავარი შედეგია. ანუ, სხვა სიტყვებით, ეს იმას ნიშნავს, რომ სადისერტაციო კვლევის ფარგლებში რიგ შემთხვევებში უკვე არსებული სისტემების გაძლიერებით, რიგ შემთხვევებში კი არსებული სისტემებიდან სრულიად დამოუკიდებლად აიგო ქართულ უნივერსალურ ჭკვიან კორპუსში ჩადგმული სახით ფუნქციონირებადი მთელი რიგი ქართული და აფხაზური ენობრივი რესურსები და ტექნოლოგიები. ესენია:

1. აფხაზური ენის თვითგანვითარებადი კორპუსი და მასში ჩადგმული აფხაზური ვებ-კლავიატურა;
2. აფხაზურ კორპუსში ჩადგმული პირველი და ჯერ-ჯერობით ერთადერთი აფხაზური მკითხველი სისტემა;
3. ჩეჩნური, ლეკური და ყაზარდოული ენების თვითგანვითარებადი კორპუსების პირველი საცდელი ვერსიები და მათში ჩადგმული ჩეჩნური, ლეკური და ყაზარდოული ვებ-კლავიატურები;
4. ქართული ტიტრირებული მეტყველების კორპუსის საცდელი ვერსია;
5. აფხაზური, ჩეჩნური, ლეკური და ყაზარდოული ტიტრირებული მეტყველების კორპუსების ამგები ინსტრუმენტები;

6. ქართულ უნივერსალურ ჭკვიან კორპუსში ჩადგმული მეგრული თვითგანვითარებადი კორპუსი;
7. მომხმარებლების მოწყობილობებიდან (კომპიუტერი, მობილური) კორპუსის გამაძლიერებელი ინსტრუმენტები;
8. ქართული და აფხაზური ტიტრირებული სამეტყველო მონაცემების სეგმენტატორების საცდელი სისტემები;
9. ქართული ჭკვიანი ფურცლის პირველი საცდელი ორენოვანი (ქართულ-აფხაზური) ვერსია (ორენოვანი ბეჭვდა; ორენოვანი კითხვა; ორენოვანი ხმოვანი მართვა; ორენოვანი ორთოგრაფიული მართლწერა; ქართული სინტაქსური მართლწერა) და მასში ჩადგმული სახით ფუნქციონირენადი:
 - 9.1. კორპუსში ჩადგმული ქართული თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსიის თითქმის სრული სამომხმარებლო ვერსია;
 - 9.2. კორპუსში ჩადგმული აფხაზური თვითგანვითარებადი ორთოგრაფიული მართლწერის საცდელ-სამომხმარებლო ვერსია;
 - 9.3. კორპუსში ჩადგმული ქართული თვითგანვითარებადი გრამატიკული მართლწერის საცდელი ვერსიის საცდელ-სამომხმარებლო ვერსია;
10. ქართული ხმოვანმართვიანი მკითხველის გაფართოების გზით შემუშავებული ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები;
 - 10.1. თანამედროვე აფხაზური მეტყველების ამოცნობისა და ხმოვანი მართვის პირველი საცდელი ვერსიები;
 - 10.2. ძველი ქართულისა და თანამედროვე აფხაზურის ხმოვანმართვიანი მკითხველების საცდელი ვერსიები;

11. ქართული N, A და V ტიპის სიტყვების ავტომატური მორფო-სინტაქსურ-სემანტიკური ანალიზატორის თითქმის სრული საცდელი და საცდელ-სამომხმარებლო ვერსიები;
12. ქართული ხმოვანი ლექსიკონის ახალი გამლიერებული ვერსია;
13. ქართული ტექსტის ამოცნობის საცდელი სისტემები;
14. აფხაზური ტექსტის ამოცნობის პირველი საცდელი სისტემები;
15. ქართული სიტყვების სრული ავტომატური ანალიზატორის (ანუ ტაგერის, დესკრიპტორის, ლემატიზატორის, გენერატორისა და ანოტატორის) პირველი საცდელი ვერსია;
16. სადისერტაციო კვლევის მთავარი შედეგი ანუ ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსია.

სადისერტაციო ნაშრომის მაღალ პროდუქტიულობას გარდა ზემოჩამოთვლილი სისტემებისა ასაბუთებს აგრეთვე კვლევის ფარგლებში ქართული ენის ლოგიკურ გრამატიკაზე დაყრდნობით შემუშავებული სრულიად ახალი თეორიული და კომპიუტერული მეთოდები, რომელთა გარეშე, ცხადია, შეუძლებელი იქნებოდა სადისერტაციო კვლევის მთავარი შედეგის ანუ თვითგანვითარებადი ქართული უნივერსალური ჭკვიანი კორპუსის პირველი საცდელ-სამომხმარებლო ვერსიისა და მასში ჩადგმული სახით ფუნქციონირებადი ზემოჩამოთვლილი უნიკალური ქართული და აფხაზური ენობრივი ტექნოლოგიური სისტემების შექმნა.

სადოქტორო კვლევის მოკლე დასკვნითი შეფასება

ამგვარად, სადოქტორო კვლევის მოკლე დასკვნითი შეფასების სახით, შეიძლება ითქვას, რომ ეს კვლევა, რომელიც ძირეულად ეყრდნობოდა საქართველოს ტექნიკური უნივერსიტეტის გრძელვადიან პროექტებს „ქართული ენის ტექნოლოგიური ანბანი“ და „აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის გეგმა-პროგრამა“, კერძოდ კი ამ გრძელვადიანი პროექტების იმ AR/122/4-105/14 ქვეპროექტს „კიდევ ერთი

ნაბიჯი მოსაუბრე ქართული თვითგანვითარებადი ინტელექტუალური კორპუსისაკენ“, რომელიც შოთა რუსთაველის საქართველოს ეროვნულმა სამეცნიერო ფონდმა 2017 წლის წარმატებულ პროექტად გამოაცხადა,²⁰ ასევე იყო წარმატებული, რადგან თამამად შეიძლება ითქვას ისიც, რომ ამ კვლევით მიღწეული შედეგებით, დღეს, უკვე კარს მომდგარ და სანახევროდ უკვე ამოქმედებულ ციფრულ ეპოქაში, ქართული და აფხაზური ენები ბევრად უფრო დაცულნი არიან ციფრული კვდომის საფრთხისგან, ვიდრე ისინი იყვნენ მანამდე. - ის, რომ ქართული ენა, მით უფრო აფხაზური, ისევე როგორც ჩეჩნური, ყაბარდოული და ლეკური ენები და, ასევე, სხვა იბერიულ-კავკასიური ანუ ქართველური და კავკასიური ენები ნაშრომის შესავალში მიმოხილულ ევროპულ ენებთან შედარებით ტექნოლოგიური მხარდაჭერის თვალსაზრისით კიდევ უკანაა, ცალსახად ასაბუთებს იმას, რომ ეს ენები დღეს მართლაც არიან ციფრული კვდომის მაღალი საფეხურის საფრთხის ქვეშ, რითაც, თავის მხრივ, ცალსახად საბუთდება წინამდებარე ნაშრომით წარმოდგენილი და ზემოთ მოკლედ მოკლედ მიმოხილული შედეგების მეტად მაღალი მნიშვნელობები როგორც სამეცნიერო (მხედველობაში გვაქვს, ხელოვნური ინტელექტისა და ბუნებრივი ენების სრული ტექნოლოგიური დამუშავების ზოგადი და მეტად აქტუალური სამეცნიერო პრობლემატიკა), ისე კულტურული (მხედველობაში გვაქვს ეკოლინგვისტური იდეოლოგიის შესაბამისად ქართული, აფხაზური და სხვა კავკასიური ენების ციფრული კვდომის საფრთხისგან დაცვის მიზნების მეტად მაღალი კულტურული მნიშვნელობები), ასევე პოლიტიკური (.მხედველობაში გვაქვს კავკასიაში ევროპული ენობრივი პოლიტიკის გატარების ანუ ენობრივი ბარიერებისგან სრულად თავისუფალი და, შესაბამისად, ეკონომიკური ურთიერთობებისთვის მაქსიმალურად გახსნილი საერთო კავკასიური კულტურული სივრცის ფორმირების მეტად მაღალი პოლიტიკური მნიშვნელობები) თვალსაზრისით [11].

²⁰ იხილეთ <https://rustaveli.org.ge/geo/tsarmatebuli-proeqebi/kidev-erti-nabiji-mosaubre-kartuli-tvitganvitarebadi-intelektualuri-korpusisaken> მისამართზე განთავსებული მასალა.

გამოყენებული ლიტერატურა

1. მალიძე შ., ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. აფხაზური ენის სრული ტექნოლოგიური უზრუნველყოფის მიზნებზე მიმართული PHDF-18-1228 პროექტის პირველი შედეგები. *საქართველოს მათემატიკოსთა კავშირის X საერთაშორისო კონფერენციის თეზისები*, 2019, 131-132.
2. მალიძე შ., ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. ქართული ენის სრული ტექნოლოგიური უზრუნველყოფის მიზნებზე მიმართული PHDF-18-1228 პროექტის პირველი შედეგები. *საქართველოს მათემატიკოსთა კავშირის X საერთაშორისო კონფერენციის თეზისები*, 2019, 132-133.
3. მალიძე შ., ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. ქართული უნივერსალური ჭკვიანი კორპუსის გაძლიერება PHDF-18-1228 პროექტით შემუშავებული მომხმარებელთა კომპიუტერების მიმართებული ახალი ინსტრუმენტებით. *საქართველოს მათემატიკოსთა კავშირის X საერთაშორისო კონფერენციის თეზისები*, 2019, 133-134.
4. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ., მალიძე შ., დემურჩევი კ. ქართული უნივერსალური ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტების საცდელი ვერსიები. *ივეკუას სახელობის გამოყენებითი მათემატიკის ინსტიტუტის გაფართოებული სემინარების თეზისები*, 2019, 29.
5. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ., მალიძე შ., დემურჩევი კ., ნ.ოქროშიაშვილი. ქართული და აფხაზური ენებით ევროკავშირში ანუ ქართული და აფხაზური ენების სრული ტექნოლოგიური უზრუნველყოფის მიზნები და პრობლემები. *საქართველოს მათემატიკოსთა კავშირის X საერთაშორისო კონფერენციის თეზისები*, 2019, 153 -154.
6. Pkhakadze K., Chikvinidze M., Chichua G., Kurtskhalia D., Malidze Sh. Georgian universal smart corpus and in it inbuilt Abkhazian, Chechen, Kabardian, Lezgian, and Migrelian self-developing corpuses results, perspectives. *Seminar of I. Vekua Institute of Applied Mathematics*, 2018, Vol. 32, 55-58.
7. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ., მალიძე შ. ევროკავშირში ტექნოლოგიურად სრულად უზრუნველყოფილი ქართული და აფხაზური ენებით. *საერთაშორისო ქართველოლოგიური კონგრესი*, 2018, 120 (ქართულად), 226 (ინგლისურად).

8. ფხაკაძე კ., ჩიჩუა გ., ჩიქვინიძე მ., კურცხალია დ., მალიძე შ. აფხაზური ხმოვანმართვიანი მკითხველი სისტემის საკომპიუტერო და საინტერნეტო ვერსიებისათვის - შედეგები და პერსპექტივები. *საქართველოს მათემატიკოსთა კავშირის IX საერთაშორისო კონფერენციის თეზისები*, 2018, 183-184.
9. ფხაკაძე კ., კურცხალია დ., ჩიქვინიძე მ., ჩიჩუა გ., მალიძე შ. აფხაზური ხმოვანი ბრაუზერისათვის - შედეგები და პერსპექტივები. *საქართველოს მათემატიკოსთა კავშირის IX საერთაშორისო კონფერენციის თეზისები*, 2018, 187-188.
10. მალიძე შ., ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. საპროექტო წინადადების - „ქართული ენით ევროკავშირში ანუ სადოქტორო თემა - ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ - საპუნლიკაციო ვერსია. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2017-2018, N11(II), 25-64.
11. მალიძე შ., ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. სადოქტორო თემის - „ქართული ჭკვიანი კორპუსის ახალი განმავითარებელი ინსტრუმენტებისა და მეთოდების შემუშავება და არსებულთა გაუმჯობესება“ - პირველი შედეგები. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2017-2018, N11(II), 65-95.
12. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ., მალიძე შ. ღია წერილი საქართველოს პარლამენტს, მთავრობას, მეცნიერებათა ეროვნულ აკადემიასა და ქართულ და აფხაზურ საზოგადოებებს ანუ სახელმწიფო ენის (ქართული, აფხაზური) სრული ტექნოლოგიური უზრუნველყოფის ერთიანი პროგრამის ძირითადი პრინციპები ანუ მომავლის კულტურულ სამყაროში ტექნოლოგიურად სრულად უზრუნველყოფილი ქართული და აფხაზური ენებით. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2017-2018, N11(II), 121-164.
13. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ., ბერიაშვილი ი., მალიძე შ. ქართული ინტელექტუალური ვებ-კორპუსი: მიზნები, მეთოდები, რეკომენდაციები. *ჟურნალის „ქართული ენა და ლოგიკა“ დამატებითი გამოცემა* (მონოგრაფია; სახელმძღვანელო). 2017, 4-320.
14. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ., მალიძე შ. ენობრივი რესურსებითა და ტექნოლოგიებით ქართული ენის სრული უზრუნველყოფის ერთიანი გეგმის პირველი კონცეპტუალური მონახაზი. *საქართველოს*

- მათემატიკოსთა კავშირის VIII საერთაშორისო კონფერენციის თეზისები*, 2017, 147-148.
15. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. მივხედოთ ქართულ ენას!. *საქართველოს მეცნიერებათა ეროვნული აკადემიის შრომები*, 2016, 149-158.
 16. Pkhakadze K., In the European Union with the Georgian Language – the Long-Term Project “Technological Alphabet of the Georgian Language” and the Threats in which is Georgian Languages. *Book of Abstracts of the VII International Joint Conference of Georgian Mathematical Union & Georgian Mechanical Union*, 2016, 192-193.
 17. ფხაკაძე კ., ჩიქვინიძე მ. ქართული ჰიბრიდული გრამატიკული (ორთოგრაფიული, სინტაქსური) მართლმწერი სისტემის პირველი საცდელ-სამომხმარებლო ვერსია. *საქართველოს მათემატიკოსთა კავშირისა და საქართველოს მექანიკოსთა კავშირის VII საერთაშორისო კონფერენციის თეზისები*, 2016, 197-198.
 18. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. ქართული ლოგიკური ანალიზატორისა და კითხვებზე მოპასუხე სისტემის პირველი ხმოვანმართვიანი საცდელი ვერსიები. *საქართველოს მათემატიკოსთა კავშირისა და საქართველოს მექანიკოსთა კავშირის VII საერთაშორისო კონფერენციის თეზისები*, 2016, 198-199.
 19. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. ქართული ჰკვიანი ფურცლის ანუ მრავალფუნქციური ქართული რედაქტორის პირველი საცდელი ვერსია. *საქართველოს მათემატიკოსთა კავშირისა და საქართველოს მექანიკოსთა კავშირის VII საერთაშორისო კონფერენციის თეზისები*, 2016, 199-200.
 20. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. ქართული ლოგიკური ამოცანებისა და ანალოგიების ავტომატურად მაწარმოებელი სისტემის პირველი საცდელ-სამომხმარებლო ვერსია. *საქართველოს მათემატიკოსთა კავშირისა და საქართველოს მექანიკოსთა კავშირის VII საერთაშორისო კონფერენციის თეზისები*, 2016, 201-202.
 21. ფხაკაძე კ., კურცხალია დ., ჩიჩუა გ., ჩიქვინიძე მ. ქართული ლოგიკური ამოცანებისა და ანალოგიების ამოხსნის უნარების ავტომატურად ტესტირების ხმოვანმართვიანი სისტემის პირველი საცდელი ვერსიები. *საქართველოს მათემატიკოსთა კავშირისა და საქართველოს მექანიკოსთა კავშირის VII საერთაშორისო კონფერენციის თეზისები*, 2016, 203-204.

22. ფხაკაძე კ., კურცხალია დ., ჩიჩუა გ., ჩიქვინიძე მ. ქართული ადაპტირებული ანუ ხმოვანმართვიანი „ჯიმეილის“ პირველი საცდელი ვერსია, *საქართველოს მათემატიკოსთა კავშირისა და საქართველოს მექანიკოსთა კავშირის VII საერთაშორისო კონფერენციის თეზისები*, 2016, 198-199.
23. ჩიჩუა გ., ფხაკაძე კ. ქართული ტიტრირებული მეტყველების სხვადასხვა სახის დამნაწევრებელი სისტემების საცდელი საინტერნეტო და არასაინტერნეტო ვერსიები. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2015-2016, N9, 209-214.
24. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ., მასხარაშვილი ა. ქართული ენის ლოგიკური გრამატიკა: საფუძვლები და გამოყენებები - ნაწილი პირველი - წინათქმა. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2015-2016, N9, 201-208.
25. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. ღია წერილი ქართულ საზოგადოებას ანუ უკვე ნამდვილად დროა ერთად მივხედოთ ქართულ ენას!. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2015-2016, N9, 131-158.
26. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. მივხედოთ ქართულ ენას!. *საქართველოს მეციერებათა ეროვნული აკადემიის ჟურნალი „მაცნე“*, 2016, 149-158.
27. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., კურცხალია დ. ქართული ენის ლოგიკური გრამატიკა: საფუძვლები და გამოყენებები - ნაწილი მესამე - ქართული ენის ლოგიკური გრამატიკის გამოყენებები. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2015-2016, N10, 4-204.
28. Pkhakadze K., Chikvinidze M., Chichua G., Kurtskhalia D., Malidze Sh. THE FIRST VERSION OF THE GEORGIAN SMART JOURNAL AND ADAPTED WIKIPEDIA. *Seminar of I. Vekua Institute of Applied Mathematics, REPORTS*, 2016, Vol. 66, 46-51.
29. ამირეზაშვილი ნ., ერემიანი რ., ლორთქიფანიძე ლ., სამსონაძე ლ., ჩიკოიძე გ., ჩუტკერაშვილი ა., ჯავახიშვილი ნ. ქართული ენის კომპიუტერული მოდელები. *საერთაშორისო კონფერენციის „ენა და თანამედროვე ტექნოლოგიები“ მასალები*, 2015, 18-21.
30. ბერიძე მ., ლორთქიფანიძე ლ., ნადარაია დ. ქართული დიალექტური კორპუსის მორფოლოგიური ანოტირების კონცეფციისათვის. *საერთაშორისო კონფერენციის „ენა და თანამედროვე ტექნოლოგიები“ მასალები*, 2015, 35-38.
31. დობორჯგინიძე ნ. ქართული ენის კორპუსის კონცეფცია. *საერთაშორისო კონფერენციის „ენა და თანამედროვე ტექნოლოგიები“ მასალები*, 2015, 57-58.

32. ლობჯანიძე ი. ქართული ენის ანალიზატორის გაუმჯობესებისა და განვითარების პერსპექტივები ქართული ენის კორპუსის საფუძველზე. *საერთაშორისო კონფერენციის „ენა და თანამედროვე ტექნოლოგიები“ მასალები*, 2015, 70-72.
33. Pkhakadze K., Chikvinidze M., Chichua G., Beriasvili I., Kurtskhalia D. The Aims and First Results of the Project “One More Step Towards Georgian Talking Self-Developing Intellectual Corpus”. *Proceedings of the international Conference on “Language and Modern Technology”*, 2015, 107-112.
34. Pkhakadze K., Chikvinidze M., Chichua G., Beriasvili I., Kurtskhalia D. IN THE EUROPEAN UNION WITH THE GEORGIAN LANGUAGE - THE AIMS AND BASEMENTS OF THE PROJECT \ONE MORE STEP TOWARDS GEORGIAN TALKING SELF-DEVELOPING INTELLECTUAL CORPUS". *Reports of Seminar of Ilia Vekua Institute of Applied Mathematics*, 2015, 37-43.
35. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., ბერიაშვილი ი., კურცხალია დ. ქართული ვებ-გვერდების ხმით მართვადი მკითხველი სისტემა. *საქართველოს მათემატიკოსთა კავშირის VI საერთაშორისო კონფერენციის თეზისები*, 2015, 168-169.
36. ჩიქვინიძე მ., ფხაკაძე კ. პროექტის „ქართული ენით ევროკავშირში ანუ სადოქტორო თემა - ქართული გრამატიკული მართლმწერი (ანალიზატორი)“ მიზნების, მეთოდებისა და პირველი შედეგების მოკლე მიმოხილვა. *საქართველოს მათემატიკოსთა კავშირის VI საერთაშორისო კონფერენციის თეზისები*, 2015, 93-94.
37. ჩიჩუა გ., ფხაკაძე კ. პროექტის „ქართული ენით ევროკავშირში ანუ სადოქტორო თემა - ქართული მეტყველების სინთეზი და ამოცნობა“ მიზნების, მეთოდებისა და პირველი შედეგების მოკლე მიმოხილვა. *საქართველოს მათემატიკოსთა კავშირის VI საერთაშორისო კონფერენციის თეზისები*, 2015, 92-93.
38. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., ბერიაშვილი ი., კურცხალია დ. ქართული ენით ევროკავშირში ანუ პროექტის „კიდევ ერთი ნაბიჯი მოსაუბრე ქართული თვითგანვითარებადი ინტელექტუალური კორპუსისაკენ“ მიზნებისა და მეთოდების მოკლე მიმოხილვა. *საქართველოს მათემატიკოსთა კავშირის VI საერთაშორისო კონფერენციის თეზისები*, 2015, 109-110.

39. Pkhakadze K. Shalva Pkhakadze's Notation Theory and Georgian Theory of the Genesis of the Different Natural Languages. *Book of Abstracts of V International Conference of Georgian Mathematical Union*, 2014, 48-49.
40. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., მასხარაშვილი ა., ფხაკაძე ნ., ბერიაშვილი ი. ქართული ენის ლოგიკური გრამატიკა როგორც ქართული ინტელექტუალური კორპუსის თეორიული საფუძველი. *საქართველოს მათემატიკოსთა კავშირის V საერთაშორისო კონფერენციის თეზისები*, 2014, 147-148.
41. Pkhakadze K., Chikvinidze M., Chichua G., Maskharashvili A., Beriasvili I. AN OVERVIEW OF THE TRIAL VERSION OF THE GEORGIAN SELF-DEVELOPING INTELLECTUAL CORPUS NECESSARY FOR CREATING GEORGIAN TEXT ANALYZER, SPEECH PROCESSING, AND AUTOMATIC TRANSLATION SYSTEMS. *Reports of Enlarged Session of the Seminar of I. Vekua Institute of Applied Mathematics*, 2014, Vol. 28, 70-75.
42. ფხაკაძე კ., ჩიჩუა გ. ქართული გამეტყველებული ტექსტების სიტყვებად, მარცვლებად და ბგერებად დამნაწევრების სხვადასხვა საცდელი სისტემები. *საქართველოს მათემატიკოსთა კავშირის V საერთაშორისო კონფერენციის თეზისები*, 2014, 136-137.
43. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ. ქართული თვითგანვითარებადი ინტელექტუალური კორპუსში ჩადგმული ქართული მეტყველების ამომცნობისა და ქართული ხმიდან-ხმაზე მთარგმნელი სისტემების საცდელი ვერსიები. *საქართველოს მათემატიკოსთა კავშირის V საერთაშორისო კონფერენციის თეზისები*, 2014, 139-140.
44. ფხაკაძე კ., ჩიქვინიძე მ. ქართულ თვითგანვითარებად ინტელექტუალურ კორპუსში ჩადგმული ქართული თვითგანვითარებადი გრამატიკული და ორთოგრაფიული მართლმწერი სისტემები. *საქართველოს მათემატიკოსთა კავშირის V საერთაშორისო კონფერენციის თეზისები*, 2014, 140-141.
45. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., მასხარაშვილი ა. ქართული ენა ციფრულ ეპოქაში და ქართული ენის ტექნოლოგიური ანბანი. *საქართველოს მათემატიკოსთა კავშირის V საერთაშორისო კონფერენციის თეზისები*, 2014, 143-144.
46. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., მასხარაშვილი ა., ფხაკაძე ნ., ბერიაშვილი ი. ქართული თვითგანვითარებადი ინტელექტუალური კორპუსის საცდელი

- ვერსია და ამ ტიპის ვებ-კორპუსის აგების მიზნები. *საქართველოს მათემატიკოსთა კავშირის V საერთაშორისო კონფერენციის თეზისები*, 2014, 144-145.
47. ფხაკაძე კ. ღია წერილი საქართველოს მეცნიერებათა ეროვნულ აკადემიას ანუ ის, რომ ევროპული ენები საფრთხის წინაშეა, სრულიად ცხადს ხდის იმ საფრთხის განსაკუთრებით მაღალ ხარისხს, რომლის წინაშეცაა ქართული! - ანუ, კვლავ ქართული ენის უფლებების დასაცავად!! - ანუ, დროა მივხედოთ ქართულ ენას!!! - მოკლე ვარიანტი. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2013-2014, N7-N8, 4-20.
48. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ., მასხარაშვილი ა., ფხაკაძე ნ., ბერიაშვილი ი. ქართული ენის ტექნოლოგიების ცენტრის ვებ-გვერდზე განთავსებული ქართული ენობრივი ტექნოლოგიური სისტემების მოკლე მიმოხილვა. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2013-2014, N7-N8, 37-64.
49. ფხაკაძე კ., ჩიქვინიძე მ., ჩიჩუა გ. პროექტი „ქართული ენის ლოგიკური გრამატიკის საფუძვლები და მისი გამოყენება საინფორმაციო ტექნოლოგიებში“ და სადოქტორო თემები „ქართული გრამატიკული მართლმწერი (ანალიზატორი)“ და „ქართული მეტყველების სინთეზი და ამოცნობა“. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2013-2014, N7-N8, 21-36.
50. Pkhakadze K., Chichua G., Chikvinidze M., Maskharashvili A. The Technological Alphabet of The Georgian Language - Aims, Methods and Results. *Reports of Enlarged Session of the Seminar of VIAM*, 2013, Vol. 27, 48-51.
51. ფხაკაძე კ. ქართული ენის ტექნოლოგიური ანბანი - XXI საუკუნის ერთ-ერთი უმთავრესი ქართული გამოწვევა. *საპარლამენტო კონფერენციის „ქართული ენა - 21-ე საუკუნის გამოწვევები“ შრომები*, 2013, 98-105.
52. მასხარაშვილი ა., ფხაკაძე კ. შემამოკლებელი გამოსახულებების ლინგვისტური კლასიფიკაცია ქართული ენის ლოგიკური გრამატიკის მიხედვით. *საქართველოს მათემატიკოსთა კავშირის IV საერთაშორისო კონფერენციის თეზისები*, 2013, 148-149.
53. ფხაკაძე კ., ჩიჩუა გ., ჩიქვინიძე მ., მასხარაშვილი ა. პროექტი “ქართული ენის ლოგიკური გრამატიკის საფუძვლები და მისი გამოყენება საინფორმაციო ტექნოლოგიებში” - სასაფუძვლო შედეგები და მისაღწევი მიზნები. *საქართველოს ტექნიკური უნივერსიტეტის 90 წლისთავისადმი მიძღვნილი*

- საერთაშორისო კონფერენციის „21-ე საუკუნის ტექნოლოგიების და განვითარების ძირითადი პარადიგმები“ შრომების კრებული, 2012, 138-146.
54. ფხაკაძე კ., ჩიჩუა გ., ჩიქვინიძე მ., მასხარაშვილი ა. ქართული ენის ლოგიკური გრამატიკის მიზნები, მეთოდები, შედეგები, *სტუ. აღიქმა გორგიძის დაბადებიდან 105 წლისთავისადმი მიძღვნილი საიუბილეო კრებული*, 2012, 129-137.
 55. Pkhakadze K., Chichua G., Chikvinidze M., Maskharashvili A. The Short Overview of the Aims, Methods, and Results of the Logical Grammar of the Georgian Language. *Reports of Enlarged Session of the Seminar of VIAM*, 2012, Vol. 26, 58-64.
 56. Taylor P. Text-to-Speech Synthesis. *Cambridge University Press*, 2009, 627.
 57. Meurer P. A Computational Grammar for Georgian, Logic, Language, and Computation. *Subseries of Lecture Notes in Computer Science, TbilLLC'2007, Springer*, 2009, 3-15.
 58. Pkhakadze K., Abzianidze L., Maskharashvili A. Mathematical analysis of Georgian Verbs. *REPORTS of Enlarged Sessions of The Seminar of I. Vekua Institute of Applied Mathematics*, 2009, Vol. 23, 87-93.
 59. Pkhakadze K., Chichua G., Abzianidze L., Maskharashvili A. 1-Stage Voice Managed Georgian Intellectual Computer System. *Reports of the Seminar of I. Vekua Institute of Applied Mathematics*, 2008, Vol.34, 96-107.
 60. Pkhakadze K., Chichua G., Abzianidze L., Maskharashvili A. Georgian Language's Theses, *Reports of the Seminar of I. Vekua Institute of Applied Mathematics*, 2008, Vol.34, 108-121.
 61. ფხაკაძე კ. წინარე ზმნური სემანტიკური ერთეულის ცნება, პირის ნიშანთა საკითხი, ინტეგრალური და არაინტეგრალური გაგების ზმნური სიტყვა-შინაარსები და ქართული ზმნის არასრული ანუ პირველი სემანტიკური კლასიფიკაცია. *ძირითადი და დამხმარე სახელმძღვანელოები თანამედროვე მათემატიკურ ენათმეცნიერებაში*, 2006, N1(II), 72-152.
 62. ფხაკაძე კ. ბუნებრივი ქართული ენობრივი სისტემის ერთი ფრაგმენტი როგორც სიმრავლურ თეორიული სისტემა. *ძირითადი და დამხმარე სახელმძღვანელოები თანამედროვე მათემატიკურ ენათმეცნიერებაში*, 2006, N1(II), 60-71.

63. ფხაკაძე კ. სიტყვათა შორის სინტაქსური კავშირების შესახებ. *ძირითადი და დამხმარე სახელმძღვანელოები თანამედროვე მათემატიკურ ენათმეცნიერებაში*, 2006, N1(II), 50-59.
64. ფხაკაძე კ., ბუაძე ნ. „ნატვრა“ ტიპის ზმნური სემანტიკები და პრინციპული განსხვავებები ქართულ და ინდოევროპულ უღლებებს შორის. *ჟურნალი „ქართული ენა და ლოგიკა“*, 2005, N2, 66-81.
65. ფხაკაძე კ. ლოგიკური ბრუნებისა და ლინგვისტური მიმართებების საკითხისათვის ქართულში. *სამეცნიერო-საგანმანათლებლო ჟურნალი „ქართული ენა და ლოგიკა“*, 2005, N1, 19-77.
66. ფხაკაძე კ., ივანიშვილი მ. ქართული ენობრივი სისტემის პირდაპირი ფორმალურ-ლოგიკური აღწერის ძირითადი მიზნებისა და მეთოდების შესახებ. *საქართველოს მეცნიერებათა აკადემიის აკად. გ.წერეთლის აღმოსავლეთმცოდნეობის ინსტიტუტის სამეცნიერო ჟურნალი „პერსპექტივა – XXI“*, 2004, 220-224.
67. ფხაკაძე კ., ჭანკვეტაძე გ., ტიბუა ლ., ივანიშვილი მ., ლეკიაშვილი ლ., სოსელია ე. წარმოადგენენ თუ არა ლოგიკური, პროგრამული და ბუნებრივი ენობრივი სისტემები ერთმანეთისაგან პრინციპულად განსხვავებული ტიპის სემანტიკურ სისტემებს?. *პარიზი – ქართულ-ევროპული ინსტიტუტის სამეცნიერო ჟურნალი „ივერია“*, 2003-2004, ტ.X-XI, 220-248.
68. ფხაკაძე კ. ბუნებრივი ქართული ენობრივი სისტემის ერთი ფრაგმენტი როგორც სიმრავლურ თეორიული სისტემა, (პირველი გამოცემა), *პარიზი – ქართულ-ევროპული ინსტიტუტის სამეცნიერო ჟურნალი „ივერია“*, 2003-2004, ტ.X-XI, 104-116.
69. Pkhakadze K., Chankvetadze G., Tibua L., Ivanishvili M., Lekiasvili, L., Soselia E. About the Main Ideas of the Direct Formal-Logical Description of the Georgian Natural Language System Through the Examples. *Proceedings of the V Tbilisi Symposium on Language, Logic and Computation, ILLC, University of Amsterdam, Tbilisi State University*, 2003, 129-137.
70. Pkhakadze K., Chankvetadze G., Tibua L., Ivanishvili M., Lekiasvili, L., Soselia E. About Main Ideas of Direct Formal-Logical Description of the Georgian Language. *Proceedings of VIAM*, 2003, Vol. 53, 33-40.

71. Pkhakadze K., Ivanishvili M. About the Main Ideas of the Direct Formal-Logical Description of the Georgian Natural Language System through One Example. *Reports of Seminar of VIAM*, 2003, Vol.23, 11-14.
72. Pkhakadze K., Chankvetadze G., Tibua L., Ivanishvili M., Lekiasvili, L., Soselia E. Logical, Programming, Natural Languages – Do They Represent Tree Different Types of the Semantic System?. *Proceedings of the I Conference of Arn.Chikobava Institute of Linguistics In “Natural Language Processing”*, 2003, 63-65.
73. უთურგაიძე თ. გრამატიკული კატეგორიებისა და მათი ურთიერთ მიმართებისათვის ქართულში. „*ქართული ენა*“, 2002, 4-244.
74. მელიქიშვილი დ. ქართული ზმნის უღლების სისტემა, 2001, 4-362.
75. ფხაკაძე კ. სიტყვათა შორის სინტაქსური კავშირების შესახებ, *პარიზი, ქართულ-ევროპული ინსტიტუტის სამეცნიერო ჟურნალი „ივერია“*, 2001, ტ.VIII-IX, 137–146.
76. ფხაკაძე კ., ივანიშვილი მ. ბუნებრივი ქართული ენობრივი სისტემის ფორმალურ-ლოგიკური იზომორფის მოძიების მიზნით. *ქართულ-ევროპული ინსტიტუტის სამეცნიერო ჟურნალი „ივერია“*, პარიზი, 2001, ტ.VI, 129-148.
77. Jurafsky D., Martin J.H. *Speech and Language Processing*. Prentice Hall, 2000, 934.
78. Hausser R. *Foundation of Computational Linguistics*. Springer, 1999, 4-534.
79. Van Benthem J., Ter Meulen A. (Editors). *Hand Book of Logic and Language*. Elsevier Science & The MIT Press, 1997, 4-1247.
80. კვაჭაძე ლ. თანამედროვე ქართული ენის სინტაქსი. 1996, 4-554.
81. Partee B.H., Meulen A., Wall E.R. *Mathematical Methods in Linguistics*. 1990, 4-663.
82. Van Benthem J. *Essays in Logical Semantics*, D. RPC, 1986, 125.
83. Harris A.C. *Georgian Syntax - A Study in Relational Grammar*. 1981, 327.
84. შანიძე ა. ქართული ენის გრამატიკის საფუძვლები. *თსუ*, 1980,1-671.
85. გამყრელიძე თ. ზმნის „პირიანობა“ და „ვალენტობა“. 1979, 33-50.
86. ჩიქობავა ა. მარტივი წინადადების პრობლემა ქართულში. *სმპ*, 1961, 4-295.
87. Пхакадзе Ш.С. Некоторые Вопросы Теории Обозначений, *ТГУ*, 1977, 4-195.