

საქართველოს ტექნიკური უნივერსიტეტი

ქართული ენის ტექნოლოგიების

სასწავლო-სამეცნიერო ცენტრი

პროექტი № 103

„ნაბეჭდი ტექსტის ქართულად გახმოვანება

და

II საფეხურის ქართული სინტაქსური მართლმწერი“

ნომინაცია: გამოყენებითი სახის მეცნიერთევადი სამეცნიერო პროექტი,
რომელსაც გააჩნია სამრეწველო რეალიზაციის უახლოესი პერსპექტივა.

პროექტის ავტორები:

ალექსანდრე ვაშალომიძე

კონსტანტინე ფხაკაძე

ტარიელ ფუტყარაძე

იანვარი, 2012 წელი

პროექტის: „ნაბეჭდი ტექსტის ქართულად გახმოვანება და II საფეხურის ქართული სინტაქსური მართლმწერი“ განხორციელება დაიწყო 2011 წლის 1 მაისს და დასრულდა 2012 წლის იანვარში. საქართველოს პოლიტექნიკური უნივერსიტეტის მიერ 2011 წლის მაის-დეკემბერში პროექტის 6 მონაწილისათვის გამოყოფილი იქნა სულ 12 500 ლარი /ხელფასის სახით/ და 3 ოფისის კომპიუტერი და 3 „ნოუთბუქი“. აქვე აღვნიშნავთ, რომ კომპიუტერები მივიღეთ ნოემბრის თვეში, რამაც გარკვეულწილად შეაფერხა ჩვენს მიერ ტექნიკურად განსახორციელებელი სამუშაოები; თუმცა, ამას ხელი არ შეუშლია პროექტის წარმატებით განხორციელებაში.

პროექტით განსაზღვრული იყო სამი ძირითადი მიზანი:

1. ქართული ენის სამეტყველო ერთეულთა ხმოვანი კომპიუტერული ბაზის შექმნით ნაბეჭდი ტექსტის ქართულად გახმოვანების მექანიზმის დამუშავება;
2. ქართული ენის ლოგიკური გრამატიკის საკითხების შემდგომი დამუშავებით ქართული გრამატიკული მართლმწერის I საფეხურის საცდელი სისტემის გაფართოება ქართული სინტაქსური მართლმწერის II საფეხურის საცდელ სისტემამდე;
3. ტექსტების კორპუსისა და ქართული ენის ტექნოლოგიების ცენტრის ვებგვერდის შექმნა.

შესაბამისად, ანგარიშს წარმოვადგენთ სამ ნაწილად.

1. ქართული ენის სამეტყველო ერთეულთა ხმოვანი კომპიუტერული ბაზის შექმნა და ნაბეჭდი ტექსტის ქართულად გახმოვანების მექანიზმის დამუშავება

სამეტყველო ერთეულთა ხმოვანი ბაზის შესაქმნელად განსახორციელებელი იყო ორი ძირითადი ამოცანა:

- სამეტყველო ერთეულთა რაოდენობის დადგენა /განსაზღვრა/ ქართული ენის ფონემატური წესების გათვალისწინებით;
- სამეტყველო ერთეულების ჩაწერა, გამოცალკევება და კომპიუტერული ბაზის შექმნა;

კერძოდ, წინა წლებში შექმნილ წარმომთქმელებში (ამძრავებში) გამოყენებული იყო მხოლოდ მარჯვენა მარტივ მარცვლოვან სეგმენტთა ხმოვანი ბაზა და ამ ტიპის არასრული ბაზებისთვის შემუშავებული მარტივი სინთეზის წესები (ესენია: Geotextreader-2009 (გ.ჩიჩუა, კ.ფხაკაძე), DemoGeotextreader-2009 (გ.ჩიჩუა, კ.ფხაკაძე) და

Geospeaker-2009 (ა.ვაშალომიძე, კ.ფხაკაძე); მათ ჰქონდათ მისაღები გარჩევადობა, მაგრამ დაბალი ნატურალურობის ხარისხი, წარმოდგენილ პროექტში გადაწყდა:

1. გასრულებულიყო ამ სითეზატორების მარჯვენა მარტივი მარცვლოვანი ხმოვანი ბაზები სრულ მარტივ მარცვლოვან ხმოვან ბაზებამდე და ქართული სამეტყველო ენის ფონემატური და არამარტო ფონემატური წესების გათვალისწინებით შემუშავებულიყო სრულ მარტივ მარცვლოვან ბაზაზე დამყარებული ბუნებრივი სინთეზის ანუ ბუნებრივი კითხვის წესები (ამ ამოცანით დაკავდნენ კ.ფხაკაძე და გ.ჩიჩუა);
2. ალოფონური მიდგომებისა და ქართული ენის ფონემატური წესების გათვალისწინებით დადგენილიყო ალოფონურ სამეტყველო ერთეულთა ხმოვანი ბაზა და ალოფონთა სინტაგმატური წესების გამოვლენის მიხედვით დაგხვეწილიყო სინთეზის წესები (ამ ამოცანით დაკავდნენ ტ.ფუტკარაძე და ა.ვაშალომიძე). ვფიქრობდით, რომ ასეთ შემთხვევაში დაიხვეწებოდა სინთეზის ხარისხი, როგორც გარჩევადობის, ასევე ბუნებრიობის თვალსაზრისით.

მარტივი მარცვლოვანი ბაზის გასრულებაზე და ამ ტიპის ბაზაზე დამყარებული მკითხველი სისტემის აგებაზე მუშაობდნენ კონსატნტინე ფაკაძე და გიორგი ჩიჩუა; ალოფონთა ბაზის შესაქმნელად და ამ ბაზის მიხედვით მეტყველების სინთეზზე მუშაობდნენ ალექსანდრე ვაშალომიძე და ტარიელ ფუტკარაძე.

პირველმა ჯგუფმა /კ. ფხაკაძე, გ. ჩიჩუა/ დაასრულა პროექტით განსაზღვრული სამუშაოები, რის შედეგადაც არა მარტო გადაჭრა პროექტით მათ წინაშე დასახული ყველა ამოცანა, ასევე, დამატებით, პროექტის ფარგლებში წარმოადგინა DemoGeotextreader_2011 სისტემა. კერძოდ:

1. ქართული სამეტყველო ენის ბუნებრივ თავისებურებათა გათვალისწინებით დაზუსტდა მარტივი ანუ არაშედგენილი სამეტყველო (სასმენო-სამეტყველო) ერთეულის ცნება და ამ მარტივი სამეტყველო ერთეულებით სხვადასხვა სახის რთული ანუ შედგენილი სამეტყველო მონაცემების ამგები მეთოდები;
2. შეიქმნა სრული ბუნებრივი მარტივი მარცვლოვანი ბაზის ჩაწერისა და გამართვის შედარებით მარტივი მექანიზმი და, ასევე, ამ მექანიზმის გამოყენებით გ. ჩიჩუას ბუნებრივი სასმენო-სამეტყველო ერთეულების სხვადასხვა ხარისხით ჩაწერილი და სხვადასხვა ხარისხით გამართული 3 ბაზა;
3. Geotextreader_2009 სისტემაზე სრული ბუნებრივი მარტივი მარცვლოვანი ბაზისა და არამორფოლოგიური მარცვლოვანი კითხვის ალგორითმის მიერთებით შეიქმნა პროექტით დაგეგმილი Geotextreader_2011 სისტემა, რომელიც თითქმის სრული გარჩევადობით, მაგრამ ჯერ კიდევ არასაკმარისი ბუნებრიობით კითხულობს ნებისმიერ ქართულ ტექსტს;

4. DemoGeotextreader_2009 სისტემებზე ბუნებრივი მარტივი მარცვლოვანი ბაზის, არამორფოლოგიური მარცვლოვანი კითხვისა და კითხვის ხმად მომხმარებლის ხმის ჩადების შედარებით მარტივი საშუალებების მიერთებით შეიქმნა DemoGeotextreader_2011 სისტემა, რომელიც საშუალებას იძლევა მასში ტექსტის ასაკრეფად გამოყოფილ ფანჯარაში აკრეფილი ტექსტის აკრეფის პარარელურად კითხვისა, რითაც ეს სისტემა უკვე დღესვე შეიძლება იქნეს გამოყენებული დაბადებით ყრუ მუნჯებისა და ამა თუ იმ მიზეზით მეტყველება უნარ წართმეული/შეზღუდული პირების საზოგადოებაში თავისუფალი ინტეგრირების მიზნით.

ამასთან, უნდა აღვნიშნოთ, რომ ჯგუფის წევრთა აზრით (კ. ფხაკაძე, გ. ჩიჩუა), მარტივი მარცვლოვანი ბაზის მიხედვით ქართული მკითხველის აგების არსებული შესაძლებლობების მხოლოდ 5-10% პროცენტია რეალიზებული. გამომდინარე აქედან, მიგვაჩნია, რომ საჭიროა შემდგომი უფორ გადრმავებული კვლევების წარმოება ამ მიმართულებით.

პირველ დანართის პირველი ნაწილით წარმოდგენილია ამ ჯგუფის მიერ შესრულებული სამუშაოს ვრცელი ანგარიში (კ.ფხაკაძე, გ.ჩიჩუა).

მეორე ჯგუფმა /ა.ვაშალომიძე, ტ. ფუტკარაძე/ დაასრულა პროექტით გათვალისწინებული დაგეგმილი სამუშაოები, კერძოდ:

- ფონემათა პარადიგმატული მახასიათებლებისა და სინტაგმატური წესების გათვალისწინებით ჩამოაყალიბა ალოფონთა ზღვრული რაოდენობის გამოყოფისთვის აუცილებელი დებულებები /კერძოდ, ალოფონები დაჯგუფდა არტიკულაციური მიდგომით: ცალკე ჯგუფებად წარმოდგენილი ხმოვნებისა და სონორების გარდა 6 ჯგუფად დაიყო ჩქამიერი ფონემები; კერძოდ, ჰომოგენური ფონემები დაჯგუფდა ლოკალური რიგების მიხედვით, მაგ., ბაგისმიერი და უკანაენისმიერი /ე.წ. ლაბიოველარული ბგერები/ მჟღერები ერთ ჯგუფში გაერთიანდნენ, წინაენისმიერები - მეორეში; შესაბამისად, ორ-ორ ჯგუფებად დაყვავით ყრუ-ფშვინვიერი და ყრუ-მკვეთრი ფონემებიც/.
- მოამზადა იმ სიტყვათა ბაზა, რომელიც შეიცავდა საჭირო ალოფონებს,
- საპატრიარქოსთან არსებულ „კოლხური კულტურის ცენტრის“ სტუდიაში /ხელმძღვანელი კომპოზიტორი მერაბ მამულაშვილი/ მსახიობი ქალის - ფატი თურმანიძის - ხმით ჩაწერა ყველა აუცილებელი სიტყვა,
 - სიტყვებიდან გამოაცალკევდა ყველა ფონემის ალოფონები (3224), დაიწყო ალოფონური ბაზის ფორმირება...

მეორე დანართში ა. ვაშალომიძის მიერ წარმოდგენილია მეორე ჯგუფის მიერ შესრულებული სამუშაოს ვრცელი ანგარიში.

ამ ეტაპზე ალოფონური ბაზის მიხედვით სინთრაქტორის შექმნა დასრულებული არ არის, მაგრამ საშუალება გვაქვს წარმოვიდგინოთ მეტყველების სინთეზი არსებული ფონემური ბაზის მიხედვით, რომელიც, ვფიქრობთ, საინტერესოა.

აქვე აღვნიშნავთ, რომ სრულყოფილი შედეგის მისაღებად აუცილებელია არა მხოლოდ ამ ბაზის დასრულება /საჭიროა კიდევ რამდენიმე თვე/, არამედ, ხმათა სრული კომპლექტის ჩაწერა, რაც გულისხმობს ექვსი დამოუკიდებელი ხმოვანი ბაზის (ასაკოვანი მამაკაცის, ასაკოვანი ქალის, ახალგაზრდა მამაკაცის, ახალგაზრდა ქალის, ბიჭის და გოგონას) შექმნას.

2. ქართული ენის ლოგიკური გრამატიკის საკითხების შემდგომი დამუშავება და ქართული გრამატიკული მართლმწერის I საფეხურის საცდელი სისტემის გაფართოება ქართული სინტაქსური მართლმწერის II საფეხურის საცდელ სისტემამდე.

პროექტში, ქართული ენის ლოგიკური გრამატიკის საკითხების დამუშავებაზე და ქართული გრამატიკული მართლმწერის I საფეხურის საცდელი სისტემის ქართული სინტაქსური მართლმწერის II საფეხურის საცდელ სისტემამდე გაფართოებაზე მუშაობდნენ კონსტანტინე ფხაკაძე და მერაბ ჩიქვინიძე.

ჯგუფმა /კ.ფხაკაძე, მ.ჩიქვინიძე/ დაასრულა პროექტით განსაზღვრული სამუშაოები. ამასთანავე, დამატებით მასთან, განსხვავებით ძველი სისტემისაგან – GeoSyntaxAnalyzer/Translator_2009, ახალი სისტემა - GeoNetSyntaxAnalyzer_2011 შესრულებულია ვებ ვარიანტად, ASP.Net-ის გამოყენებით, რითაც გაუმჯობესდა პროექტით დაგეგმილი შედეგი.

ჯგუფის წევრებს მიაჩნიათ (კ. ფხაკაძე, მ. ჩიქვინიძე) და ჩვენც ვვარაუდობთ, რომ ეს არის, ქართული ენის ლოგიკური გრამატიკის რთული და მარტივი თხრობითი წინადადებებით აგებული ტექსტების სინტაქსურად შემმოწმებელი პირველი ექსპერიმენტული სისტემა ქართულისათვის. ამგვარად, პროექტის ფარგლებში წარმოდგენილია:

1. ქართული ენის ბუნებრივი თავისებურებებისა და ქართული ენის ლოგიკური გრამატიკის ფარგლებში უკვე არსებული ქართული ენის პირველი საფეხურის მათემატიკური თეორიის (ე.ი. ქართული ენის მარტივი თხრობითი წინადადებების მათემატიკური თეორიის) საფუძველზე წინასწარი სახით შემუშავდა ქართული ენის მეორე საფეხურის მათემატიკური თეორია (ე.ი. ქართული ენის რთული თხრობითი წინადადებების მათემატიკური თეორია);

2. პროექტით დაგეგმილის სახით, მაგრამ ვებ ვერსიაში, შეიქმნა GeoNetSintaxAnalizer_2011 სისტემა, რომელიც არის ქართული ენის რთული და მარტივი თხრობითი წინადადებების სინტაქსური მართლწერის შემმოწმებელი ექსპერიმენტული სისტემა.

აქვე უნდა აღვნიშნოთ, რომ მნიშვნელოვანია ამ მიმართულებით გაფართოვდეს კვლევითი და დანერგვითი პროცესები.

პირველი დანართის მეორე ნაწილით წარმოდგენილია ამ ჯგუფის მიერ შესრულებული ამ სამუშაოს ვრცელი ანგარიში (კ.ფხაკაძე, მ.ჩიქვინიძე).

3. ქართული ენის ტექნოლოგიების ცენტრის ვებგვერდისა და ტექსტების კორპუსის შექმნა.

ამ სამუშაოს შესრულება ეთხოვა როლანდ მათიაშვილს. მის მიერ შესრულებული სამუშაო იხ., დანართში /3/, ხოლო პროდუქცია მისამართზე: <http://www.gtu.ge/gltc/index.htm>

პროექტის ხელმძღვანელი:

ტექნიკურ მეცნიერებათა

დოქტორი

ალექსანდრე ვაშალომიძე

დანართი I - დანართი შედგენილია კონსტანტინე ფხაკაძის მიერ

საქართველოს ტექნიკური უნივერსიტეტი

საანგარიშო მოხსენება
პროექტზე № 103

ნაბეჭდი ტექსტის ქართულად გახმოვანება
და

II საფეხურის ქართული სინტაქსური მართლმწერი

საქართველოს ტექნიკური უნივერსიტეტის
ქართული ენის ტექნოლოგიზების სასწავლო-სამეცნიერო
ცენტრი

„ჩვენ უნდა შევქმნათ კონკურენტუნარიანი საინფორმაციო გარემო, რომელიც ზემოაღნიშნულ ფასეულობებს დაეფუძნება და მათ თანამედროვეობასთან შეაზავებს. ეს კი დიდად დაეხმარება ახალგაზრდებს, დამოუკიდებლად გასცეს პასუხი დროის გამოწვევებს.

ამაშია საქართველოს გადარჩენა.“

სრულიად საქართველოს კათალიკოს-პატრიარქი,
 მცხეთა-თბილისის მთავარ ეპისკოპოსი,
 ბიჭვინთისა და ცხუმ-აფხაზეთის მიტროპოლიტი ილია II

შობა ქრისტესი
 თბილისი, 2012 წელი

შ ი ნ ა ა რ ს ი

| | | |
|---|---|----|
| 1 | პროექტის ფარგლებში ჩატარებული სამუშაოებისა და მიღწეული შედეგების მოკლე კვარტალური ანგარიში | 9 |
| 2 | პირველი საპროექტო მიმართულებით პროექტის ფარგლებში ჩატარებული სამუშაოებისა და მიღწეული შედეგების შემაჯამებელი ანგარიში | 12 |
| 3 | მეორე საპროექტო მიმართულებით პროექტის ფარგლებში ჩატარებული სამუშაოებისა და მიღწეული შედეგების შემაჯამებელი ანგარიში | 18 |
| 4 | ბოლოთქმის მაგიერ | 50 |
| 5 | დამოწმებული ლიტერატურის ნუსხა | 51 |

1 პროექტის ფარგლებში ჩატარებული სამუშაოებისა და მიღწეული შედეგების მოკლე კვარტალური ანგარიში

№ 103 პროექტის ავტორთა ჯგუფმა (ტ.ფუტკარაძე, ა.ვაშალომიძე, ვ.ფხაკაძე) პროექტის ძირითად კვლევით მიზნად დასახა ქართული ენის სამეტყველო ერთეულთა ხმოვანი ბაზის შექმნისა და ქართული ენის ლოგიკური გრამატიკის საკითხების დამუშავების გზით ქართული ნაბეჭდი ტექსტის მკითხველი მექანიზმისა და ქართული გრამატიკული მართლმწერის II საფეხურის საცდელი სისტემის აგება. შესაბამისად, პროექტზე მომუშავე ჯგუფი სამუშაოებს წარმართავდა შემდეგი ორი ძირითადი მიმართულებით:

1. ქართული ენის სამეტყველო ერთეულთა ხმოვანი ბაზისა და ნაბეჭდი ტექსტის ქართულად გახმოვანების მექანიზმის შექმნა;
2. ქართული ენის ლოგიკური გრამატიკის საკითხების დამუშავება და ქართული გრამატიკული მართლმწერის I საფეხურის საცდელი სისტემის გაფართოება ქართული სინტაქსური მართლმწერის II საფეხურის საცდელ სისტემამდე.

ამასთან, პროექტის ძირითადი მიზნების შესაბამისად, პროექტის ძირითადი ამოცანები იყო:

1. ქართული ენის სამეტყველო ერთეულთა ხმოვანი ბაზის შექმნა;
2. ქართული ენის ლოგიკური გრამატიკის საკითხების დამუშავება;
3. ნაბეჭდი ტექსტის ქართულად გახმოვანების მექანიზმის შექმნა;
4. ქართული გრამატიკული მართლმწერის I საფეხურის საცდელი სისტემის გაფართოება ქართული სინტაქსური მართლმწერის II საფეხურის საცდელ სისტემამდე.

თავიდანვე უნდა აღინიშნოს, რომ პროექტზე მუშაობისას ჩვენმა ქვეჯგუფმა (ვ.ფხაკაძე, გ.ჩიჩუა, მ.ჩიქვინიძე) არამხოლოდ თავი გაართვა ყველა ზემოთ დასახულ მიზანსა და ამოცანას, არამედ, უფრო მეტიც, ჯგუფი ყველა ამ საპროექტო მიზნისა და ამოცანის გასწვრივ გასცდა დაგეგმილ მიჯნებს და დამატებით შეასრულა მნიშვნელოვანი ხასიათის სამუშაოები. ამასთან, პროექტის ძირითადი მიზნებისა და ამოცანების მიხედვით პროექტის ფარგლებში ჩვენს მიერ შესრულებულ სამუშაოთა მოკლე კვარტალური განაწერი შემდეგია:

I საპროექტო მიმართულება: ქართული ენის სამეტყველო ერთეულთა ხმოვანი ბაზისა და ნაბეჭდი ტექსტის ქართულად გახმოვანების მექანიზმის შექმნა (ვ.ფხაკაძე, გ.ჩიჩუა):

- **II კვარტალი -**
 - ქართული სამეტყველო ენის ბუნებრივ თავისებურებათა გათვალისწინებით დაზუსტდა მარტივი ანუ არაშედეგნილი სამეტყველო (სასმენო-სამეტყველო) ერთეულის ცნება და ამ მარტივი სამეტყველო ერთეულებით სხვადასხვა სახის რთული ანუ შედეგნილი სამეტყველო მონაცემების ამგები მეთოდები;
 - შემუშავდა Geotextreader_2009 და DemoGeotextreader_2009 სისტემებზე ბუნებრივი სამეტყველო ბაზის (ანუ სრული მარტივი მარცვლოვანი ბაზის) მიერთების გზები და ასეთი ბაზით ტექსტის კითხვის ალგორითმი,

რომელსაც ჩვენ ბუნებრივი არამორფოლოგიური მარცვლოვანი კითხვის ალგორითმს ვუწოდებთ.

• **III კვარტალი -**

- Geotextreader_2009 და DemoGeotextreader_2009 სისტემებზე ბუნებრივი მარტივი მარცვლოვანი ბაზის და არამორფოლოგიური მარცვლოვანი კითხვის მიერთებით შეიქმნა Geotextreader_2011 და DemoGeotextreader_2011 სისტემების წინასწარი ვერსიები;
- Geotextreader_2011 და DemoGeotextreader_2011 სისტემების გასრულების მიზნით დაიწყო მუშაობა მათში ინტეგრირებული ბუნებრივი მარტივი მარცვლოვანი ბაზის შესაქმნელად. - ეს სამუშაოები ჩატარდა იმ დაბალი ხარისხის ტექნიკური უზრუნველყოფის პირობებში, რაც ჯგუფს ახალი ტექნიკის მიღებამდე ჰქონდა.

• **IV კვარტალი -**

- დასრულდა Geotextreader_2011 და DemoGeotextreader_2011 სისტემებში ინტეგრირებული ბუნებრივი მარტივი მარცვლოვანი ბაზის შექმნის პროცესი. ამასთან, შეიქმნა ამ ბაზის ის ცალი, რომელიც გ.ჩიჩუას ხმითაა შევსებული (ამ ბაზას, პირობითად, 8-ე ბაზას ვუწოდებთ);
- პროექტით გათვალისწინებული სახით აიგო Geotextreader_2011 სისტემა. ამასთან, დამატებით, აიგო DemoGeotextreader_2011 სისტემა, რომელიც აღიჭურვა მასში მომხმარებლის ხმის საბაზო მონაცემების მარტივად ჩამწერი და გამმართავი დანამატით;
- ახალი ტექნიკის მიღების შემდეგ საპატრიარქოსთან არსებულ „კოლხური კულტურის ცენტრის“ სტუდიაში (ხელმძღვანელი - კომპოზიტორი მერაბ მამულაშვილი) მაღალი ხარისხის მიკროფონზე გ.ჩიჩუას ხმით გაკეთდა მარტივი მარცვლოვანი ბაზის მონაცემთა ჩაწერა, რომელიც დღეს უკვე ნაწილობრივ გამართულია. - ჩანაწერის ჟღერადობა ეჭვქვეშ არ აყენებს მის მაღალ ხარისხს, თუმცა, ამ შემთხვევაში, ცუდი ის იყო, რომ ჩაწერის მომენტში ჩვენ არ გვქონდა DemoGeotextreader-2011 სისტემაზე მიერთებული ხმის ჩამწერი და გამმართავი დანამატებით სარგებლობის საშუალება, რამაც, საბოლოო ჯამში, გაართულა ამ მაღალხარისხიანი ჩანაწერის გამართვის პროცესი (ამ ბაზას, პირობითად, საუკეთესოს ვუწოდებთ);
- ახალი ტექნიკის მიღების შემდეგ შეიქმნა გ.ჩიჩუას ხმოვან ერთეულთა კიდევ ერთი ბაზა (ამ ბაზას, პირობითად, 9-ე ბაზას ვუწოდებთ), რომელიც გამართვის პროცესშია;
- წინა წლებში (იხ. [6], [9], [11], [12], [13], [21], [23]) და პროექტის ფარგლებში წარმოებული კვლევების შეჯერებით მომზადდა თეორიული ნაშრომი „ბუნებრივი ქართული სამეტყველო ერთეულის ცნება და ქართული მრავალხმოვანი მკითხველი სისტემის ამგები მეთოდებისა და აგების მიზნების მიმოხილვა“ (იხ. საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის შრომები, N1, თბილისი, 2011, გვ.31-51);

II საპროექტო მიმართულება: ქართული ენის ლოგიკური გრამატიკის საკითხების დამუშავება და ქართული გრამატიკული მართლმწერის I საფეხურის საცდელი

სისტემის გაფართოება ქართული სინტაქსური მართლმწერის II საფეხურის საცდელ სისტემამდე (კ.ფხაკაძე, მ.ჩიქვინიძე):

- **II კვარტალი -**

- ქართული ენის ლოგიკური გრამატიკის ფარგლებში უკვე არსებული ქართული ენის I საფეხურის მათემატიკური თეორიისა (იხ. [30]), და ქართული ენის ლოგიკური გრამატიკის შემუშავების მიზნით წინა წლებში და პროექტის ფარგლებში წარმოებული კვლევების შედეგად (იხ. [38], [36], [34], [33], [28], [25], [24], [22], [18], [17], [16], [9], [5], [4], [2], [1]) ნაწილობრივი სახით შემუშავდა ქართული ენის II საფეხურის მათემატიკური თეორია. კერძოდ, მონტეგეიუსეული ზოგადი მათემატიკური მიდგომებითა (იხ. [50], [45]) და აღნიშვნათა თეორიის საფუძველზე (იხ. [49], [44], [41], [7]) ქართული ენის ლოგიკურ გრამატიკაში შემუშავებული მეთოდებით (იხ. [38], [35], [34], [32], [29], [26], [20], [14]) დაკლასიფიცირდა და მათემატიკურად აღიწერა:
 - ქართული ენის რთული თხრობითი წინადადებების ამგები სიტყვები;
 - მძიმის სასვენ ნიშნად წოდებული ‘,‘ სიმბოლოს სხვადასხვა ტიპის ყოფაქცევები ქართული ენის რთულ წინადადებებში;
 - ქართული ენის რთული თხრობითი წინადადებები;
- შემუშავდა GeoSintaxAnalyzer/Translator_2009 სისტემის GeoSintaxAnalyzer_2009 ქვესისტემის ASP.NET ფორმატში GeoNetSintaxAnalyzer_2011 სისტემად გადაწერის კონცეფცია, რათა ნაცვლად პროექტით გათვალისწინებული მართლმწერის C პროგრამული სისტემისა შეგვექმნა მართლმწერის ვებ სისტემა.

- **III კვარტალი -**

- მათემატიკური შესწავლის საფუძველზე, ქართული ენის რთული წინადადებების ტიპობრივი მრავალფეროვნების გათვალისწინებით (იხ. [43], [32], [16]), გამოიყო ქართული ენის რთული წინადადებების ძირეული ნაწილი და ნაწილობრივი სახით შემუშავდა ამ ტიპის რთული წინადადებების მათემატიკური თეორია;
- GeoSintaxAnalyzer/Translator_2009 სისტემის GeoSintaxAnalyzer_2009 ქვესისტემის ASP.NET ფორმატში გადაწერით შეიქმნა GeoNetSintaxAnalyzer_2011 სისტემის პირველი წინასწარი ვერსია;
- ქართული ზმნისა და სახელის ბუნებრივი ლოგიკური თავისებურებების გათვალისწინებით:
 - გაფართოვდა ქართული ენის ლოგიკური გრამატიკის ფარგლებში შემუშავებული ზმნისა და ზმნის უღლების ბუნებრივი მათემატიკური მოდელები. კერძოდ, გამოიყო 0_ადგილიანი, 1_ადგილიანი, 2_ადგილიანი და 3_ადგილიანი უღლებები და უღლებათა თითოეული ამ ტიპისათვის აღიწერა ქართული ენის უღლებათა (პარადიგმათა) მაწარმოებელი წესები;
 - გაფართოვდა ქართული ენის ლოგიკური გრამატიკის ფარგლებში შემუშავებული სახელისა და სახელის ლოგიკური ბრუნების ბუნებრივი მათემატიკური მოდელები. კერძოდ, დამუშავდა ლოგიკური ბრუნების ა_გავრცობილი და ც_გავრცობილი

ვარიანტები და კვლევის ამ ეტაპზე სახელთა ეს ვარიანტული ბრუნვითი ფორმები ტიპობრივად გაიგივდა მათ ძირითად ცალკეობთან;

- გაფართოვდა GeoNetSyntaxAnalyzer_2011 სისტემის უკვე არსებული წინასწარი ვერსია. კერძოდ:
 - სისტემას მიუერთდა ფორმაუცვლელ სიტყვათა ბაზა;
 - სისტემა გაფართოვდა სხვადასხვა ტიპის უღლებისა და უღლების წესების მაწარმოებელი კომპიუტერული ინსტრუმენტებით, რომლებიც იძლევიან ახალი ზმნებით სისტემის ავტომატური გაფართოების ფართო საშუალებებს;
 - სისტემა გაფართოვდა სახელთა ა_ვარიანტული და ც_ვარიანტული ბრუნვებთან მაწარმოებელი კომპიუტერული ინსტრუმენტებით, რომლებიც იძლევიან ახალი სახელებით სისტემის ავტომატური გაფართოების ფართო საშუალებებს.

• IV კვარტალი -

- წინასწარი სახით შემუშავდა ქართული ენის მეორე საფეხურის მათემატიკურ თეორია ანუ მომზადდა ნაშრომი „ქართული მარტივი და რთული წინადადებების ლოგიკური სინტაქსისა და ქართული მართლმწერის ამგები მეთოდების და აგების მიზნების მიმოხილვა“ (იხ. საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის შრომები, N1, თბილისი, 2011, გვ.52-92);
- პროექტით დაგეგმილის სახით, მაგრამ ვებ ვერსიაში, შეიქმნა GeoNetSyntaxAnalyzer_2011 სისტემა, რომელიც არის ქართული ენის რთული და მარტივი თხრობითი წინადადებების სინტაქსური მართლმწერის შემმოწმებელი პირველი ექსპერიმენტული სისტემა;

2 პირველი საპროექტო მიმართულებით პროექტის ფარგლებში ჩატარებული სამუშაოებისა და მიღწეული შედეგების შემაჯამებელი ანგარიში

I საპროექტო მიმართულებით ჯგუფს ვალდებულებად ჰქონდა აღებული ქართული ენის სამეტყველო ერთეულთა ხმოვანი ბაზისა და ქართული ნაბეჭდი ტექსტის მკითხველი მექანიზმის შექმნა. ამასთან დაკავშირებით, მოგახსენებთ:

ქართული მეტყველების სინთეზატორის შექმნაზე მუშაობა საქართველოში დიდი ხანია მიმდინარეობს. შედეგად გვაქვს სხვადასხვა ავტორებისა და ჯგუფების მიერ შექმნილი სხვადასხვა სინთეზატორები.¹ მათ შორის, თავად ჩვენ ჯგუფში, ჯერ

¹ სამწუხაროდ, არცერთი ეს სინთეზატორი არ აკმაყოფილებს იმ საერთაშორისოდ არსებულ სტანდარტებს, რაც დღეს ამ ტექნოლოგიურ სექტორშია. მათ შორის არც ის სინთეზატორი, რომელიც 2009 წელს დამზადდა საქართველოს უსინათლოთა საზოგადოების დაკვეთით მინსკში მაღალი ტექნოლოგიების ბელორუსული ჯგუფის მიერ. ამ სინთეზატორთან დაკავშირებით მომზადებულ მინსკის ტელევიზიის რეპორტაჟში (სრულად იხ. <http://www.gtu.ge/gltc/links/video.htm> მისამართზე)

კიდევ 2009 წლამდე, უკვე იყო რამდენიმე განსხვავებული მიდგომებით (იხ. [40], [23], [13], [12], [11]) აგებული და განსხვავებული ხარისხით მომუშავე სინთეზატორი.²

2009 წელს, თსუ სახელმწიფო მიზნობრივი პროგრამის "კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში" 9 თვიანი ქვეპროგრამის "ქართული ენის და აზროვნების მათემატიკური თეორიის საფუძვლები და ქართული ინტელექტუალური კომპიუტერული სისტემის პირველი ვერსიის ნაწილები" (პროექტის ხელმძღვანელი კ.ფხაკაძე) ფარგლებში შემუშავდა ჩვენს მიერ ე.წ. მარტივი მარცვლოვანი მიდგომებით აგებული და მომხმარებლის ხმის ჩამდები საშუალებებით აღჭურვილი ორი სხვადასხვა გვარად და სხვადასხვა ხარისხით მომუშავე სინთეზატორი (იხ. [11], [12], [10]).³ ესენია, Geotextreader_2009 (გ.ჩიჩუა,

ამკარად გადაჭარბებულად ფასდება მისი მნიშვნელობა. კერძოდ, რეპორტაჟში არაერთხელ გაესვა ხაზი იმას, რომ მსოფლიოში პირველად ქართული სინთეზატორი 2010 წელს დაამზადდა ამ ბელორუსული ჯგუფის მერ იმ სახელშეკრულებო კონტრაქტის ფარგლებში, რომელიც მათთან 2010 წლის გაზაფხულზე საქართველოს უსინათლოთა კავშირმა გააფორმა, რაც, რბილად რომ ვთქვათ, არ შეესაბამება სინამდვილეს.

გარდა ამისა, რეპორტაჟში, ხაზი ესმება იმასაც, რომ მათ ეს კონტრაქტი მეტად დაბალ - დამკვეთისათვის სარფიან ფასად შეასრულეს. არადა, რამდენადაც ჩვენთვის ცნობილია, ეს ცხრათვიანი კონტრაქტი საქართველოს უსინათლოთა საზოგადოებას 200 000 ლარზე მეტი დაუჯდა! - აქ არ შეიძლება გულის ტკივილით არ აღინიშნოს, რომ ამ 200 000 ლარის საქართველოში დახარჯვის შემთხვევაში, დღეს უკვე, დამაკმაყოფილებელი ხარისხის ქართული სინთეზატორით ისარგებლებდა არამარტო საქართველოს უსინათლოთა საზოგადოება, არამედ მთელი ქართული საზოგადოება! - ამ განაცხადის ამკარა დასტურია ჩვენი ჯგუფის მიერ მარტივი მარცვლოვანი ბაზის საფუძველზე შემუშავებული ხმის ჩამდები საშუალებებით აღჭურვილი სინთეზატორები.

² <http://gllc.ge/news.html> და <http://www.gtu.ge/gllc/links/video.htm> მისამართებზე შეგიძლიათ ნახოთ თსუ სახელმწიფო მიზნობრივი პროგრამის "კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში" კვლევითი მიზნების გათვალისწინებით 2006-2009 წლებში ჩვენ ჯგუფში მიზერული დაფინანსებისა და მიზერული ტექნიკური უზრუნველყოფის პირობებში შემუშავებული სინთეზატორების საცდელი ვერსიების მიერ სინთეზირებული ხმების ნიმუშები.

³ ქართულ ენაში მეტყველების სინთეზის მარტივ მარცვლოვან ბაზაზე დამყარების იდეა ეკუთვნის კ.ფხაკაძეს და ეს იდეა გამოიკვეთა 2004-2009 წლებში ბუნებრივი ქართული მეტყველების შესწავლის მიზნით თსუ სახელმწიფო მიზნობრივი პროგრამის "კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში" ფარგლებში წარმოებული კვლევების შედეგად. ასევე, ამავე წლებში, კ.ფხაკაძის ხელმძღვანელობით შემუშავდა მომხმარებლის ხმის ჩამდები საშუალებებით აღჭურვილი სინთეზატორის კონსტრუირების მეთოდოლოგია, როგორც მეტყველების სინთეზირებისა და ამოცნობის მაღალი ხარისხით მომუშავე სისტემების აგების ერთადერთი გზა. - ამასთან, AgoraNet-ის ვებ-გვერდზე (<http://www.modeltalker.com/>), განთავსებულია მათ მიერვე შემუშავებული ModelTalker სისტემის წინასწარი ვერსია, რომელსაც ისინი მეტყველების მასინთეზებელ რევოლუციურ პროგრამულ სისტემას უწოდებ, რადგან მათი ეს სისტემა მეტყველება უნარ შეზღუდულ ადამიანებს (როგორც მათ ვინც ახლა კარგავს ხმას, ასევე მათ ვისაც უკვე დაეკარგა ეს უნარი) საშუალებას აძლევს თავიანთივე ხმის სინთეზირებული ვერსიებით დაამყაროს თავისუფალი სამეტყველო ურთიერთობა საზოგადოების სხვა წევრებთან. კერძოდ, მათივე თქმით „**The ModelTalker System is a revolutionary speech synthesis software package designed to benefit people who are losing or who have already lost their ability to speak. It allows people with ALS or other conditions to use a synthetic version of their own voice for communication, or to choose a voice best suited to represent them**“. - არადა, როგორც მოგახსენეთ, ჩვენ, მომხმარებლის ხმის ჩამდები ინსტრუმენტებით აღჭურვილი სინთეზატორის აგებაზე მუშაობა 2005 წლიდან დავიწყეთ და 2007-2009 წლებში უკვე გვექონდა ასეთი სისტემების არაუარყოფითი რეალიზაციები (იხ. [9], [10], [11], [12], [13]).

კ.ფხაკაძე), DemoGeotextreader_2009 (გ.ჩიჩუა, კ.ფხაკაძე) და Geospeaker_2009 (ა.ვაშალომიძე, კ.ფხაკაძე).⁴

ამგვარად, ნაბეჭდი ტექსტის ქართულად გახმოვანების მექანიზმი, რისი აგებაც პროექტით ჯგუფს პასუხისმგებლობად ჰქონდა აღებული, არის Geotextreader_2011 (გ.ჩიჩუა, კ.ფხაკაძე), რომელიც Geotextreader_2009 სინთეზატორზე დაყრდნობით და მისი შემდგომი გაფართოებით შემუშავდა პროექტის ფარგლებში წარმოებული კვლევების შედეგად. ასევე, პროექტის ფარგლებში წარმოებული კვლევების შედეგად დამატებით შემუშავდა DemoGeotextreader_2011 (გ.ჩიჩუა, კ.ფხაკაძე), რომელიც, თავის მხრივ, DemoGeotextreader_2009 სისტემის შემდგომი გაფართოების შედეგია და რომელიც მნიშვნელოვანია იმით, რომ იგი აღჭურვილია მომხმარებლის ხმის ჩამდები საშუალებებით, რაც სისტემას დამატებით მეტად მნიშვნელოვან პერსპექტივებს სძენს.

ქვემოთ, სტრუქტურული და ფუნქციონალური თვალსაზრისებით, მოკლედ და დეტალურად აღვწერთ Geotextreader_2011 და DemoGeotextreader_2011 სისტემებსა და მათი შესაძლო გამოყენებათა არეებს. ამასთან, ის მეთოდები, რომლებიც ამ სისტემების აგებისას იქნა გამოყენებული და რომლებიც ქართული სამეტყველო ენის სპეციფიკების გათვალისწინებით შემუშავდა, დეტალურადაა მიმოხილული ნაშრომში „ბუნებრივი ქართული სამეტყველო ერთეულის ცნება და ქართული მრავალხმოვანი მკითხველი სისტემის ამგები მეთოდებისა და აგების მიზნების მიმოხილვა“ (კ.ფხაკაძე, გ.ჩიჩუა, საქართველოს ტექნიკური უნივერსიტეტის სასწავლო-სამეცნიერო ცენტრის შრომები, N1, თბილისი, 2011 გვ.31-51).

2.1. Geotextreader_2011 და DemoGeotextreader_2011 სისტემებისა და მათი შესაძლო გამოყენებათა არეების მიმოხილვა.

Geotextreader_2011 სისტემა კითხულობს ვორდის დოკუმენტს, აღიქვამს როგორც ლათინური კოდირების, ასევე უნიკოდში ჩაწერილ ტექსტებს. პროგრამის მუშაობისთვის გახსნილი უნდა იყოს ვორდის დოკუმენტი. პროგრამა კითხულობს აქტიურ დოკუმენტს. კითხვის დაწყებისათვის უნდა დააჭიროთ **Speak** ღილაკს. შესაძლოა კითხვის გაგრძელება დოკუმენტის ნებისმიერი ადგილიდან, ამისათვის მაუსი უნდა დააწკაპუნოთ დოკუმენტის შესაბამის პოზიციაში. შესაძლოა ხმის

⁴ ამ სინთეზატორების ხმის ნიმუშები შეგიძლიათ მოისმინოთ <http://www.gtu.ge/gltc/index.htm> მისამართზე. - ერთ-ერთი ძირითადი მიზეზი იმისა, რომ ქართული მკითხველი სისტემების კითხვის ხარისხი ჯერ კიდევ ვერ უტოლდება ამ სფეროში არსებულ საერთაშორისო სტანდარტებს, არის ის, რომ ამ სექტორში დაწინაურებულ სხვა სამეტყველო ენებთან შედარებით ქართული სამეტყველო ენა არასაკმარისად არის შესწავლილი.

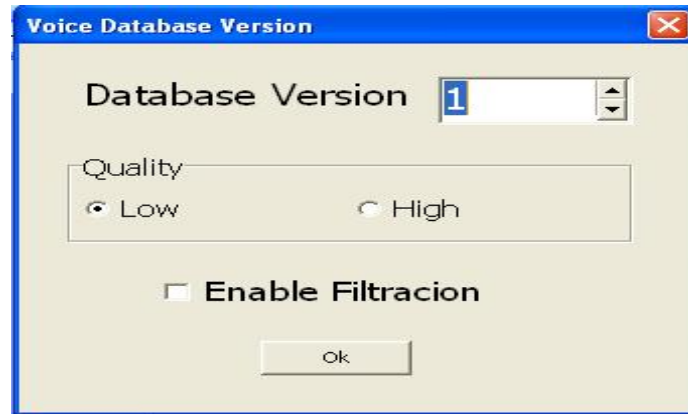
აქ მხედველობაში გვაქვს როგორც მარტივი სიტყვის შიგნით არსებული ფონეტიკური, ასევე ენის მორფოლოგიასა და სინტაქსზე მიბმული რთული ენობრივი კონსტრუქციების სემანტიკური სამეტყველო სპეციფიკები. - ასეთი კვლევების გარეშე, ცხადია, შეუძლებელია ქართულისათვის აიგოს გააზრებული კითხვის უნარის მქონე სისტემა.

ამასთან, თავად მიზეზი იმისა, რომ ჯერ კიდევ ქართული სამეტყველო ენა მისი სრული და ამომწურავი ტექნოლოგიების მიზნის გათვალისწინებით არასაკმარისად არის შესწავლილი, არის ამ საკითხისადმი ადგილობრივ არსებული არასწორი დამოკიდებულება, რისი ამკარა დასტურია 2006 წლამდე თსუ-ში მათემატიკისა და კომპიუტერულ მეცნიერებათა მიმართულებით არსებული ლინგვისტური სასწავლო ხაზების გაუქმება, ნაცვლად ამ სასწავლო ხაზების იმ დროს მკაფიოდ გამოკვეთილ მაღალ მოთხოვნათა შესაბამისი მოდერნიზაციისა.

შეცვლა, ამისათვის პროგრამის ფორმაზე უნდა დააწკაპუნოთ მაუსის მარჯვენა ღილაკი და გამოსულ მენიუში აირჩიოთ მანიშნებელი **Voice**. ეკრანზე გამოვა დიალოგი (იხ., ქვემოთ, სურ. 1). დიალოგში შეიძლება შეცვალოთ:

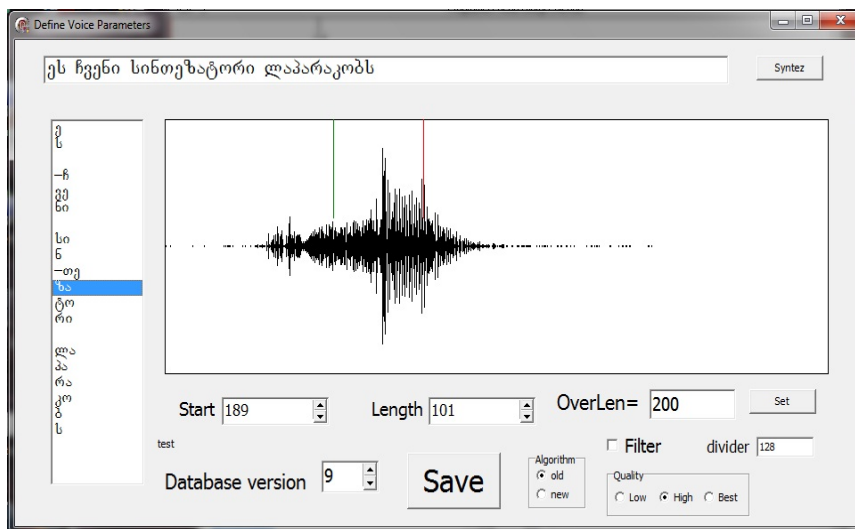
1. **Database Version** ღილაკის მეშვეობით ხმა ანუ მონაცემთა ბაზის ნომერი, რაც უზრუნველყოფს მომხმარებლის სასურველი ხმით კითხვას;

2. **Quality** ბლოკში **Low** (შეესაბამება 11025 ჰც) და **High** (შეესაბამება 44100 ჰც) ღილაკების მეშვეობით შეგიძლიათ შეცვალოთ კითხვის სიხშირე.



სურ. 1

DemoGeotextreader_2011 სისტემა კითხულობს სპეციალურად ამისთვის გამოყოფილ ფანჯარაში ქართულად აკრეფილ ნებისმიერ ტექსტს. ასევე, სისტემა მომხმარებელს საშუალებას აძლევს მასთან დიალოგის რეჟიმში სისტემაში ჩადოს თავისი ან სხვა რომელიმე მისთვის სასურველი ადამიანის ხმა და გამართოს ანუ დაარეგულიროს იგი, რის შემდეგაც ამ უკვე ჩადებული ხმით სისტემა მომხმარებელს წაუკითხვას ნებისმიერ ტექსტს.



სურ. 2

DemoGeotextreader_2011 სისტემასთან მუშაობისას ეკრანზე (იხ. სურ. 2) გამოდის პროგრამასთან დიალოგის დასამყარებელი ფორმა, რომელზეც განთავსებულია:

1. **Syntez დილაკი**, რომელზეც მიბმულ ფანჯარაში მომხმარებელს შეუძლია აკრიფოს ნებისმიერი ტექსტი, რის შემდეგაც დილაკის გააქტიურებით სისტემა უკვე შერჩეული ხმით წაიკითხავს აკრიფებულ ტექსტს და, ამდროულად, სისტემის ფორმაზე მარცხნივ განთავსებულ ვერტიკალურ ფანჯარაში მოგვცემს წასაკითხი ტექსტის დამარცვლულ სახეს.

ეს მომხმარებელს საშუალებას აძლევს საჭიროების შემთხვევაში ხმის სარეგულირებელი ჰორიზონტალური ფანჯრისა და მასზე მიბმული დილაკების დახმარებით გაასწოროს (დაარეგულიროს) საბაზო მონაცემი;

2. **Database Version დილაკი**, რომელზეც მიბმულია ფანჯარა, რომლის დახმარებითაც მომხმარებელს შეუძლია აირჩიოს ის ხმა, ანუ ის მონაცემთა ბაზა, რომლითაც მას სურს სისტემამ წაიკითხოს ტექსტი;

3. **Quality მანიშნებელი**, რომელზეც მიბმულია ბლოკი, რომლის დილაკები **Low** (შეესაბამება 11025 ჰც) და **High** (შეესაბამება 44100 ჰც) მომხმარებელს საშუალებას აძლევს აირჩიოს სიხშირე;

4. **OverLen დილაკი**, რომელზეც მიბმულია ღია ფანჯარა, რომლის დახმარებითაც მომხმარებელი ირჩევს საბაზო ხმების გადაფარვის სიღრმეს, რითაც საბოლოო ჯამში კითხვის სიჩქარე რეგულირდება;

5. **Filter დილაკი**, რომელზეც მიბმულია ფანჯარა, რითაც მომხმარებელს შეუძლია ჩართოს ან გამორთოს ფილტრი;

6. **Divider დილაკი**, რომელზეც მიბმულია ფანჯარა, რომლითაც მომხმარებელს 16 ბიტის ჩანაწერის შემთხვევაში შეუძლია უზრუნველყოს სახმო მონაცემის ფანჯარაში ხმის გრაფიკის განთავსება;

7. **Start** და **Length** დილაკები, რომელზეც მიბმული ფანჯრებით მომხმარებელს მარტივი საშუალება ეძლევა თავად დაარეგულიროს ხმის სარეგულაციო ფანჯარაში გრაფიკის სახით გამოსული მონაცემი;

8. **Save დილაკი**, რომლითაც მომხმარებელი უკვე დარეგულირებულ (გამართულ) სახმო მონაცემს ინახავს ბაზაში;

9. **Algorithm მანიშნებელი**, რომელზეც მიბმულია ბლოკი, რომლის დილაკებით **old** (შეესაბამება კითხვის ანუ სინთეზის ძველ ალგორითმს) და **new** (შეესაბამება კითხვის ანუ სინთეზის ახალ ალგორითმს) მომხმარებელს საშუალება ეძლევა აირჩიოს კითხვის ძველი/ახალი რეჟიმი.

ამგვარად, პროექტის შესრულების შედეგად, დასრულებული პროდუქტის სახით შეიქმნა Geotextreader_2011 და DemoGeotextreader_2011 სისტემები, თუმცა, ეს არ უნდა იყოს გაგებულნი ისე თითქოს ეს სისტემები მათი ამგები მეთოდოლოგიით განსაზღვრული შესაძლებლობების ზღვრულ რეალიზაციებს წარმოადგენს! - პირიქით, ვალდებულნი ვართ ხაზი გავუსვათ, რომ ამ სისტემებში ჯერ ჩადებულია (ე.ი.

ათვისებულა) მათი ამგები მეთოდოლოგიის მხოლოდ 5-10%, რაც იმის გათვალისწინებით, რომ შედეგი მაინც თვალშისაცემია, არაუმნიშვნელო ფაქტია.

ამასთან, Geotextreader_2011 და DemoGeotextreader_2011 სისტემების გამოყენება, გამომდინარე იქიდან, რომ ბოლომდე გამართულ ბაზაზე (როგორცაა, მაგალითად, 8-ე ბაზა) ისინი იძლევიან თითქმის სრულად გარჩევადი და საკმარისად ბუნებრივი კითხვის ხარისხს, უკვე შესაძლებელია მათი როგორც DVD დისკზე კოპირების, ასევე მათი საინტერნეტო ფორმატში რეალიზაციის გზით.

თუმცა, ამ ეტაპზე, ჩვენ უფრო უპრიანად მივიჩნევთ ტექნიკური უნივერსიტეტის ვებ-გვერდზე ამ სისტემების ოდნავ მოდიფიცირებული ვერსიების საინტერნეტო რეალიზაციების განთავსებას ფართო მოხმარების მიზნით. კერძოდ, ამ სისტემების მომხმარებლებად შეიძლება იქცნენ:

1. **ბიბლიოთეკები**, რათა, ერთი მხრივ, არაუსინათლოებს შესთავაზონ ელექტრონული ტექსტებიდან ინფორმაციის მიღების დამატებითი სერვისი, მეორე მხრივ, მოემსახურონ უსინათლოებს;
2. **საინფორმაცია საშუალებები**, რომლებიც რიგ შემთხვევებში უდიეტოროდ - მხოლოდ ეკრანზე გამოტანილი ტექსტობრივი სახით აწვდიან ინფორმაციას მომხმარებლებს;
3. **ნებისმიერი კერძო პირი**, ვინც ფლობს და ამა თუ იმ დატვირთვით მოიხმარს ქართულ ენას ან სწავლობს მას იმ მიზნით, რომ შემდგომში ამა თუ იმ დატვირთვით მოიხმაროს იგი;
4. **განათლებისა და მეცნიერების სამინისტრო**, რათა ახალ საფეხურზე აიყვანოს და, შესაბამისად, ხარისხობრივად გააუმჯობესოს ინტერნეტში მოზარდებისათვის სასწავლო-საგანმანათლებლო პროგრამების ინტენსიური და მიზნობრივი მიწოდების უკვე დაწყებული პროცესი;
5. **დაბადებით ყრუ-მუნჯები ან ამა თუ იმ მიზეზით მეტყველების უნარ შეზღუდულები/წართმეულები**. ამ თვალსაზრისით განსაკუთრებით საინტერესოა DemoGeotextreader-2011 სისტემა,⁵ რომელსაც აქვს სპეციალური ფანჯარა, რომელშიც აკრეფილი ტექსტი შესაძლებელია წაკითხულ (გამეტყველებულ) იქნას აკრეფის პარალელურად. ამგვარად, ამ სისტემის საფუძველზე მარტივ შესაძლებელია⁶ გაკეთდეს ქართულკლავიატურიანი მობილური ტელეფონებისათვის ისეთი მათშივე ინტეგრირებული ან მათზე მიერთებადი დანამატი, რომლის მეშვეობითაც ზემოაღნიშნული კატეგორიის პირი ამ ტელეფონს სხვა პირებთან სრულფასოვან სამეტყველო ურთიერთობაში შესვლის

⁵ <http://www.modeltalker.com/comparison.html> მისამართზე, AgoraNet-ის ვებ-გვერდზე, განთავსებულია მათი სისტემით სინთეზირებული ხმები და ამ ხმების ბუნებრივი ნიმუშები, რომლებიც გარკვეულ შემთხვევებში გარკვეული ხარისხით გვანან მათ სინთეზირებულ ვარიანტებს. თუმცა, თუ შევადარებთ ჩვენთან ამავე მიმართულებით წინა წლებში და პროექტის ფარგლებში მიღწეულ შედეგებს, დაგვერწმუნებით, რომ ჩვენი შედეგები AgoraNet-ის შედეგებს არ ჩამოუვარდება არც სინთეზის, და არც ჩადებული ხმის სინთეზირებულ ხმასთან მიმსგავსების თვალსაზრისით! - და ეს მიუხედავად იმისა, 2007 წლიდან ჩვენ ფაქტიურად ყველანაირი საშუალება წაგვერთვა თსუ სახელმწიფო-მიზნობრივი პროგრამით „კომპიუტერის სრული პროგრამულ-მომსახურეობითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ მიმდინარე კვლევების წარმოებისა! - არადა, 2007 წელს, ჩვენ უკვე ვმუშაობდით არა მხოლოდ ხმის ჩამდები ინსტრუმენტებით აღჭურვილი მკითხველი სისტემების აგებაზე, არამედ ასეთი სისტემების ხმის მართვის ბლოკით აღჭურვაზეც.

⁶ მიუხედავად იმისა, რომ ეს მართლაც მარტივ შესაძლებელია, ჩვენთვის ცხადია ისიც, რომ საქართველოს ტექნიკური უნივერსიტეტის დაუხმარებლად და მისგან დამოუკიდებლად, სტუ „ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრს“, ამის გაკეთება გაუჭირდება.

მიზნითაც გამოიყენებს (მას ამისთვის დასჭირდება მხოლოდ მობილურის კლავიატურით სათქმელი ტექსტის აკრეფა);

6. ქართული ანბანით უზრუნველყოფილი კომპიუტერების, მობილურების, აიფონების მომხმარებლები, ასევე, ის საწარმოები, რომლებიც უშვებენ ასეთ მოწყობილობებს.⁷ - ამ შემთხვევაშიც მარტივ შესაძლებელია გაკეთდეს ასეთ მოწყობილებაში ინტეგრირებული ან მასზე მიერთებადი მარტივი დანამატი, რომლის მეშვეობითაც ამ მოწყობილობების მომხმარებელი მიიღებს დამატებით მოხერხებულ საშუალებას მხედველობის დაძაბვის გარეშე - მოსმენით მიიღოს ინფორმაცია მოწყობილებაში არსებულ ქართულ ტექსტობრივ მასალაზე (იქნება ეს, ფოსტა, მესიჯი, თუ ინტერნეტიდან ცოცხლად შემომავალი თუ თავად ამ მოწყობილებაში არსებული ტექსტობრივი მონაცემი);
7. დიასპორის საკითხებში საქართველოს სახელმწიფო მინისტრის აპარატის ვებ-გვერდი „ქართული დიასპორა“. - სისტემის სასაიტე ვერსიის ამ ვებ-გვერდზე განთავსება, ამუშავება და მისი თანდათანობითი განახლება-გაუმჯობესება,⁸ ცხადია, ხელს შეუწყობს ქართული დიასპორის ქართულ ენაში და, შესაბამისად, ქართველ ერთან უფრო მჭიდრო კულტურულ-ეკონომიკური ინტეგრირების დღეს მიზნობრივად მიმდინარე პროცესს.

3 მეორე საპროექტო მიმართულებით პროექტის ფარგლებში ჩატარებული სამუშაოებისა და მიღწეული შედეგების შემაჯამებელი ანგარიში

II საპროექტო მიმართულებით ჯგუფს ვალდებულებად ჰქონდა აღებული ქართული ენის ლოგიკური გრამატიკის საკითხების დამუშავება და ქართული გრამატიკული მართლმწერის I საფეხურის საცდელი სისტემის გაფართოება ქართული სინტაქსური მართლმწერის II საფეხურის საცდელ სისტემამდე. ამასთან დაკავშირებით, მოგახსენებთ:

მანქანური თარგმნის (და არა მხოლოდ მანქანური თარგმნის) ამოცანის გათვალისწინებით ქართული სინტაქსური (გრამატიკული) მართლმწერის შექმნაზე მუშაობა საქართველოში წინა საუკუნის 60-იანი წლებიდან მიმდინარეობს.⁹

⁷ ამ თვალსაზრისით განსაკუთრებით საინტერესოა ჩვენს მიერ წარმოდგენილი სისტემების უნარი, დაიმატოს ახალი საკითხავი ხმა, რაც მომხმარებელს საშუალებას აძლევს წაკითხოს თავის მოწყობილებას ჩატვირთული მასალა საკუთარი ან სხვა ნებისმიერი მის მოწყობილებაში ასევე მის მიერ ჩადებული ხმით.

⁸ როგორც აღინიშნა, Geotextreader_2011 და DemoGeotextreader_2011 სისტემებში ჯერ ათვისებულია მათი ამგები მეთოდოლოგიის მხოლოდ 5-10%. - ამას ვამბობთ იმის გათვალისწინებით, რომ ფონეტიკოსის მონაწილეობის გარეშე, ვიწრო ჯგუფში ჩატარებული თეორიული კვლევების შედეგად, უკვე მონიშნული გვაქვს ამ სისტემების შემდგომი ხარისხობრივი გაუმჯობესების ეტაპები, რაც ფონეტიკოსის ჩართულობით, ცხადია, კიდევ უფრო სრულყოფილ სახეს მიიღებს. - ამასთან, ბუნებრივ მარტივ მარცვლოვან ბაზაზე დამყარებული სინთეზირების პლიუსი ინტეგრალურ სქემაში მისი ადვილად განთავსებასთან ერთად სტანდარტებამდე მისი ხარისხობრივი ამაღლების შედარებითი სიიაფეცაა ნებისმიერ სხვა სახის სინთეზირების მეთოდთან შედარებით!

⁹ იხ. კრებული „მანქანური თარგმანი“, რომელიც საქართველოს მეცნიერებათა აკადემიის სარედაქციო-საგამომცემლო საბჭოს დადგენილებით გამოიცა გამომცემლობაში „მეცნიერება“ (გამოცემის

მიუხედავად ამისა, დღემდე, არსებობს ქართული ენის სინტაქსური მართლწერის შემოწმებელი მხოლოდ ის ექსპერიმენტული სისტემები, რომლებიც ჩვენ ჯგუფში შემუშავდა 2003 წლიდან თსუ სახელმწიფო მიზნობრივი პროგრამის "კომპიუტერის სრული პროგრამულ-მომსახურეობითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში" ფარგლებში კ.ფხაკაძის ხელმძღვანელობით წარმოებული კვლევების შედეგად. კერძოდ, ასეთი სისტემის აგების მიმართულებით პირველი იმ დროს მნიშვნელოვანი, მაგრამ მაინც პრიმიტიული სახის ექსპერიმენტული შედეგები მიღწეულ იქნა 2001-2004 წლებში (იხ. [38], [36], [35], [34], [33], [32]).

ამ ექსპერიმენტულმა შედეგებმა სისტემური თვალსაზრისით შედარებით უფრო დასრულებული ექსპერიმენტული შედეგების სახე მიიღო 2005-2006 წლებში უნივერსიტეტებში სამეცნიერო კვლევების ხელშემწყობი პროგრამის ფარგლებში თსუ სახელმწიფო მიზნობრივ პროგრამაზე „კომპიუტერის სრული პროგრამულ-მომსახურეობითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ განათლებისა და მეცნიერების სამინისტროს მიერ გამოყოფილი სახელმწიფო საბიუჯეტო საგრანტო დაფინანსების ფარგლებში კ.ფხაკაძის ხელმძღვანელობით მიმდინარე კვლევებით (იხ. [30], [29], [28], [26], [25], [24]).¹⁰

სამწუხაროდ, მოხდა ისე, რომ 2006 წლიდან, სახელმწიფო მიზნობრივი პროგრამის "კომპიუტერის სრული პროგრამულ-მომსახურეობითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში" კვლევითი მიზნების გათვალისწინებით წინა წლებში თსუ-ში მიმდინარე სასწავლო, სამეცნიერო და სასწავლო-სამეცნიერო პროცესები თითქმის სრულად შეჩერდა.¹¹

ამან განაპირობა ის, რომ ქართული სინტაქსური მართლწერის აგების მიზნით 2004-2006 წლებში უკვე არსებული არაუმნიშვნელო შედეგების შემდგომი გაღრმავების

რედაქტორი ვ.ჭიჭინაძე) 1965 წელს. ასევე, მიხეილ ჩხაიძის ცნობილი ნაშრომი „ავტომატური თარგმანი“, რომელიც თბილისის უნივერსიტეტის გამომცემლობამ გამოსცა 1981 წელს.

¹⁰ 2005 წლის 1 ივლისიდან 31 დეკემბრამდე განათლებისა და მეცნიერების სამინისტრომ თსუ სახელმწიფო მიზნობრივ პროგრამაზე „კომპიუტერის სრული პროგრამულ-მომსახურეობითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ (პროექტის ავტორი და ხელმძღვანელი კ.ფხაკაძე) გამოყო 12 400 ლარიანი სახელმწიფო საბიუჯეტო საგრანტო დაფინანსება.

ამ დაფინანსებით, გარდა უკვე აღნიშნულისა, ქართული ენის სრული მათემატიკური შესწავლისა და ტექნოლოგიების მიმართულებით გადაიდგა რიგი მნიშვნელოვანი ნაბიჯები. მათ შორის: თსუ მათემატიკის მიმართულებით ამუშავდა „ბუნებრივი ენების ლოგიკის“ საბაკალვრო სპეციალიზაცია (ეს საქართველოსთვის ახალი სპეციალიზაცია თსუ-ში დამტკიცდა კ.ფხაკაძის მიერ წარდგენილი პროგრამის საფუძველზე); თსუ 2006 წლის „მათემატიკის“ სამაგისტრო პროგრამაში კ.ფხაკაძის მიერ წარდგენილი პროგრამით დამტკიცდა „მათემატიკური ლინგვისტიკის“ მოდული ერთ-ერთ ცალკე აღებულ სამაგისტრო სპეციალიზაციად; დაფუძნდა და გამოიცა სამეცნიერო-საგანმანათლებლო ჟურნალის „ქართული ენა და ლოგიკა“ (მთავარი რედაქტორი კ.ფხაკაძე) პირველი ორი ნომერი; დაფუძნდა და 2005 წლის 7-14 აპრილს ჩატარდა ქართული ენის დღისადმი მიძღვნილი პირველი რესპუბლიკური სასემინარო კვირეული „ლოგიკა, ენა, ინფორმატიკა“; ითარგმნა და საუნივერსიტეტო სახელმძღვანელოს სახით გამოიცა ბ.პარტის, ა.მიულენისა და რ.უოლის ცნობილი სახელმძღვანელოს - „მათემატიკური მეთოდები ლინგვისტიკაში“ - პირველი ნაწილი „ბუნებრივი და მათემატიკური ენების ზოგადი ლოგიკური საფუძვლები“ (მთარგმნელ-რედაქტორი კ.ფხაკაძე); ითარგმნა და სტუდენტებისათვის დამხმარე სახელმძღვანელოდ გამოიცა ვ.გრუნველდის კარგად ცნობილი ლექსიკონი „ლოგიკა და ენა - მცირე განმარტებითი ლექსიკონი“ (მთარგმნელ-რედაქტორი კ.ფხაკაძე); სტუდენტებისათვის ამუშავდა ღია საუნივერსიტეტო სემინარი „ლოგიკა, ენა, გამომთვლელი“ (სემინარის ხელმძღვანელი კ.ფხაკაძე);

¹¹ ის, თუ რამ განაპირობა ეს, გამომდინარე ამ სახელმწიფო მიზნობრივი პროგრამის მეტად მაღალი აქტუალობიდან და იმ შედეგებიდან, რაც ამ მეტად მაღალი ეროვნული მნიშვნელობების მქონე საკითხთა მიმართულებით (იხ. [27], [19], [10], [9], [2]) უკვე მიღწეული იყო როგორც სასწავლო, ისე სამეცნიერო თვალსაზრისებით, ჩემთვის დღესაც გამოცანად რჩება!

პროცესები 2009 წლამდე არსებითად იყო შეფერხებული. მიუხედავად ამისა, მიზნობრივ პროგრამაზე მომუშავე ჯგუფმა, 2006-2009 წლებში, შეძლო ქართული ენის შესწავლის მიზნით წინა წლებში მიმდინარე თეორიული კვლევების შენარჩუნებაც და უკვე მიღწეული ექსპერიმენტული შედეგების მცირედი გაუმჯობესებაც (იხ. [22], [21], [20], [19], [18], [17], [16], [14]).

2009 წლის 1 აპრილიდან 2009 წლის 31 დეკემბრამდე საქართველოს საპატრიარქოს წმინდა ანდრია პირველწოდებულის სახელობის ქართულმა უნივერსიტეტმა გამოყო საგრანტო დაფინანსება თსუ სახელმწიფო მიზნობრივი პროგრამის „კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ ქვეპროგრამაზე „ქართული ენის და აზროვნების მათემატიკური თეორიის საფუძვლები და ქართული ინტელექტუალური კომპიუტერული სისტემის 1-ვერსიის ნაწილები“. ¹² - ქართული სინტაქსური მართლმწერის აგების მიმართულებით პირველი არსებითი ხასიათის შედეგი მიღწეულ იქნა სწორედ ამ ცხრათვიანი პროექტის ფარგლებში. კერძოდ, ამ პროექტით, ერთ მთლიან სისტემაში ინტეგრირებული სახით შემუშავდა GeoSintaxAnalyzer/Translator_2009 სისტემა (მ.ჩიქვინიძე, კ.ფხაკაძე), რომელიც აერთიანებს შემდეგ ერთმანეთთან ფუნქციონალურად დაკავშირებულ ქვესისტემებს:

- I საფეხურის ქართული სინტაქსური მართლმწერი სისტემა;¹³
- ქართულ-მათემატიკური ორმხრივ მთარგმნელი სისტემა, რომელიც, თავის მხრივ, I საფეხურის ქართულ სინტაქსურ მართლმწერში ცალკე ქვესისტემის სახით არის ჩადგმული;
- ქართული მრავალენოვანი (ინგლისური, გერმანული) ორმხრივ მთარგმნელი სისტემა;

ამგვარად, II საფეხურის ქართული სინტაქსური მართლმწერის საცდელი (ექსპერიმენტული) სისტემა, რისი აგებაც ქართული ენის ლოგიკური გრამატიკის საკითხების კვლევისა I საფეხურის ქართული სინტაქსური მართლმწერის საცდელი სისტემის გაფართოების გზით პროექტით ჯგუფს პასუხისმგებლობად ჰქონდა აღებული, არის GeoNetSintaxAnalyzer_2011 სისტემა (მ.ჩიქვინიძე, კ.ფხაკაძე),¹⁴ რომელიც ქართული ენის ლოგიკური გრამატიკის ძირეული ტიპის რთული და მარტივი

¹² პროექტი, რომლის ხელმძღვანელი იყო კ.ფხაკაძე და რომელშიც 5 ძირითად მკვლევართან ერთად 3 სტუდენტი და ერთი გერმანული ენის სპეციალისტიც იყო ჩართული, დაფინანსდა 19 375 ლარით.

¹³ ადგილობრივ არსებული ვითარების მკაფიოდ წარმოჩენის მიზნით ხაზგასასმელია, რომ GeoSintaxAnalyzer/Translator_2009 სისტემაში ინტეგრირებული I საფეხურის ქართული სინტაქსური მართლმწერი, რომელიც წინა წლებში თსუ სახელმწიფო მიზნობრივი პროგრამის „კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ ფარგლებში კ.ფხაკაძის მიერ ქართული ენის მათემატიკური თეორიისა და ქართული ენის ლოგიკური გრამატიკის შემუშავების მიზნით წარმოებული თეორიული კვლევების შედეგებს ემყარება, არის ქართული ენის მარტივი და მარტივ განვრცობილი წინადადებების სინტაქსურად მაანალიზებელი და შემმოწმებელი პირველი ექსპერიმენტული პროგრამული სისტემა. - ჩვენი ეს განაცხადი ემსახურება იმ მეტად სავალალო ვითარების მძაფრად წარმოჩენის მიზანს, რომლის წინაშედაც დღეს ამ კუთხით ქართული ენა იმყოფება! - <http://ena.ge/georgian-grammar-electronic-system-v1.0> მისამართზე, შეგიძლიათ ნახოთ ინტერნეტში განთავსებული ალტერნატიული ქართული პროგრამული პროდუქტი. - ეს უფრო სრულ სურათს შექმნის აქ ხაზგასმულ ვითარებაზე. - აქ უნდა აღინიშნოს <http://spelling.ge/index.php> მისამართზე განთავსებული ქართული ორთოგრაფიული მართლმწერიც, რომელიც ძლიერ დამეხმარა ტექსტის გამართვაში.

¹⁴ განთავსებულია <http://www.gtu.ge/gltc/index.htm> და <http://translate.dmind.net/Default.aspx> მისამართზე.

წინადადებებით აგებული ტექსტების სინტაქსურად მანალიზებელი და შემსწორებელი პირველი ექსპერიმენტული პროგრამული სისტემა ქართულისათვის.

ამასთან, ამ საპროექტო მიზნის შემთხვევაშიც, ჯგუფი გასცდა პროექტით დაგეგმილ მიზნას, და ნაცვლად II საფეხურის ქართული სინტაქსური მართლმწერის ჩვეულებრივი C პროგრამული ვერსიისა ჩასაბარებლად წარმოადგინა სისტემის ASP.NET ფორმატში დამზადებული ვებ ვერსია, რაც ქართული ენის სინტაქსური მართლმწერის აგების მიმართულებით პროექტის ფარგლებში წინ გადადგმულ მნიშვნელოვან ნაბიჯს დამატებით ღირებულებას სძენს მისი გამოყენებითი არიალის ამომწურავი გაფართოებისა და სრული გამარტივების თვალსაზრისით.

ქვემოთ, სტრუქტურული და ფუნქციონალური თვალსაზრისებით, მოკლედ და დეტალურად აღწერთ GeoNetSintaxAnalizer_2011 სისტემას, ამ სისტემის შესაძლო გამოყენებათა არეებსა და განვითარებათა პერსპექტივებს. ამასთან, ის მეთოდები, რომლებიც ამ სისტემების აგებისას იქნა გამოყენებული და რომლებიც ქართული ენის სპეციფიკების გათვალისწინებით ქართული ენის მეორე საფეხურის მათემატიკური თეორიის ფარგლებში შემუშავდა, დეტალურადაა მიმოხილული ნაშრომში „ქართული მარტივი და რთული წინადადებების ლოგიკური სინტაქსი და ქართული მართლმწერი სისტემის ამგები მეთოდებისა და აგების მიზნების მიმოხილვა“ (კ.ფხაკაძე, მ.ჩიქვინიძე), საქართველოს ტექნიკური უნივერსიტეტის სასწავლო-სამეცნიერო ცენტრის შრომები, N1, თბილისი, 2011, გვ.52-92).

3.1. GeoNetSintaxAnalizer_2011 სისტემა და მისი შესაძლო გამოყენებათა არეებისა და განვითარებათა პერსპექტივების მიმოხილვა.

GeoNetSintaxAnalizer_2011 არის GeoSintaxAnalizer_2009 სისტემის შემდგომი გაფართოების შედეგი, რომელიც, თავის მხრივ, GeoSintaxAnalizer/Translator_2009 სისტემის ერთ-ერთი ძირეული ქვესისტემაა. ამგვარად, GeoNetSintaxAnalizer_2011 სისტემის აღწერამდე ძალიან მოკლედ აღწერთ GeoSintaxAnalizer/Translator_2009 სისტემას.

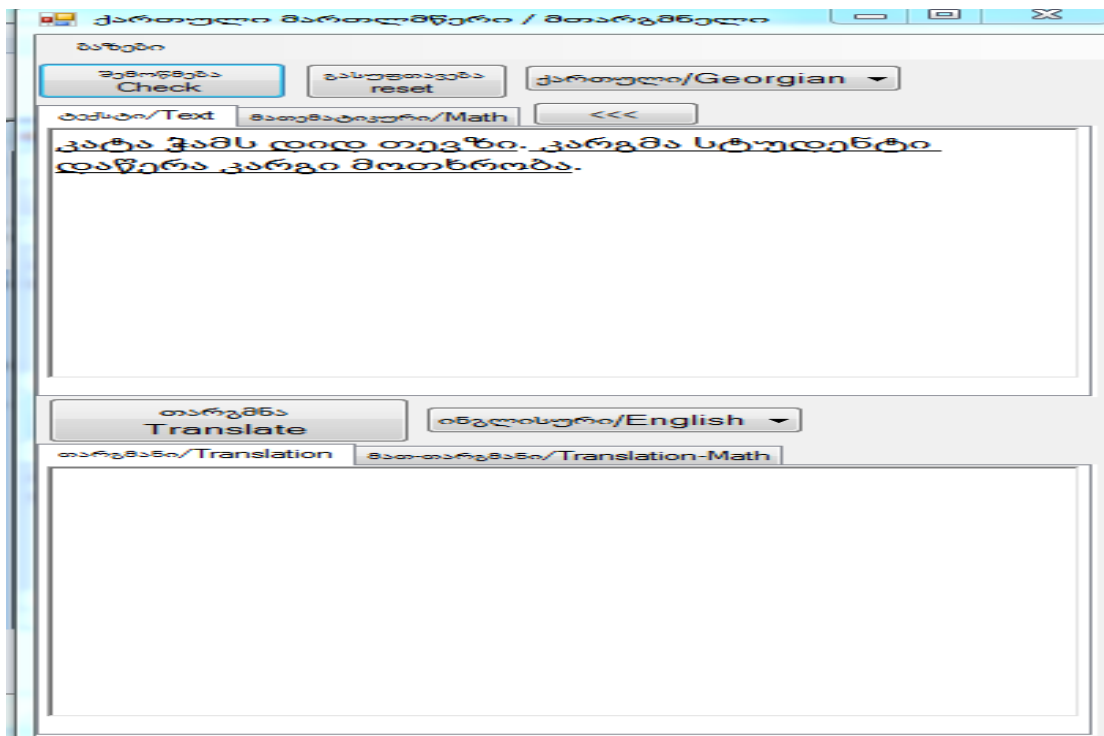
ამასთან, GeoNetSintaxAnalizer_2011 სისტემის აღწერამდე, მოკლედ აღწერთ აგრეთვე ModusPonens_2005 (კ.ფხაკაძე, კ.ფხაკაძე) სისტემას, რომელიც ორპირობიანი ქართული ლოგიკური ამოცანების ამომხსნელი და შემმოწმებელი პირველი ქართული ექსპერიმენტული სისტემაა (იხ. [26]) და რომელიც 2005 წელს თსუ სახელმწიფო მიზნობრივი პროგრამის „კომპიუტერის სრული პროგრამულ-მომსახურეობითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ ფარგლებში კ.ფხაკაძის მიერ ქართული ინტელექტუალური კომპიუტერული სისტემის აგების მიზნით წარმოებული კვლევებით შემუშავებული მიდგომებით მომზადდა ექსელის გარემოში.

GeoSintaxAnalizer/Translator_2009 და ModusPonens_2005 სისტემების ეს აღწერები, მკითხველს დაეხმარება როგორც GeoNetSintaxAnalizer_2011 სისტემის მეტად ფართო პერსპექტიული მნიშვნელობების განჭვრეტაში, ასევე იმ სამუშაოების შეფასებაში, რაც ჯგუფმა პროექტის ფარგლებში ამ სისტემის ასაგებად შეასრულა.

3.1.1. GeoSintaxAnalizer/Translator_2009 სისტემის აღწერაში დაგვეხმარება სისტემის გარე ფორმა (იხ., ქვემოთ, სურ. 3). სისტემას აქვს ორი ფანჯარა (პირობითად - ზედა და ქვედა). სინტაქსურად გასაანალიზებელი/სათარგმნი ტექსტი იწერება ზედა ფანჯარაში. ამასთან:

1. იმისათვის, რომ ტექსტი სინტაქსურად გაანალიზდეს (შემოწმდეს) უნდა გააქტიურდეს ზედა ფანჯრის **შემოწმება** ლილაკი, რის შედეგადაც სისტემა წინადადებას, რომელშიც შეცდომაა დაშვებული, უკეთებს ხაზგასმას;¹⁵
2. თარგმნისას ენათა წყვილი ირჩევა ზედა და ქვედა ფანჯრების ენის გადამრთველების მეშვეობით (სურ. 3-ზე ზედა ველში ეს გადამრთველი დაყენებულია - **ქართული/Georgian**, ქვედა ველში კი - **ინგლისური/English** პოზიციაზე);
3. მას შემდეგ, რაც ენათა წყვილი არჩეულია ქვედა ველის **თარგმნა/Translate** ლილაკის გააქტიურება ქვედა ველში იძლევა თარგმნილ ტექსტს;
4. ზედა ველში **ტექსტი/Text** და **მათემატიკური/Math**, ქვედა ველში კი - **თარგმანი/Translation** და **მათ-თარგმანი/Translation-Math** ლილაკების გააქტიურება უზრუნველყოფს ამ ველებში განთავსებული ენობრივი/მათემატიკური ტექსტების მათემატიკურ/ენობრივ თარგმნებს;
5. ლილაკი **გასუფთავება/reset** იძლევა ზედა ველში ჩაწერილი ტექსტის დამუშავების დამხმარე საშუალებებს;
6. ლილაკი **ბაზები** იძლევა წვდომას სისტემის ქართულ, ინგლისურ და გერმანულ ბაზებზე, რომელთაგან ქართულზე საუბარი გვექნება პროექტით შემუშავებული GeoNetSyntaxAnalyzer_2011 სისტემის აღწერისას.

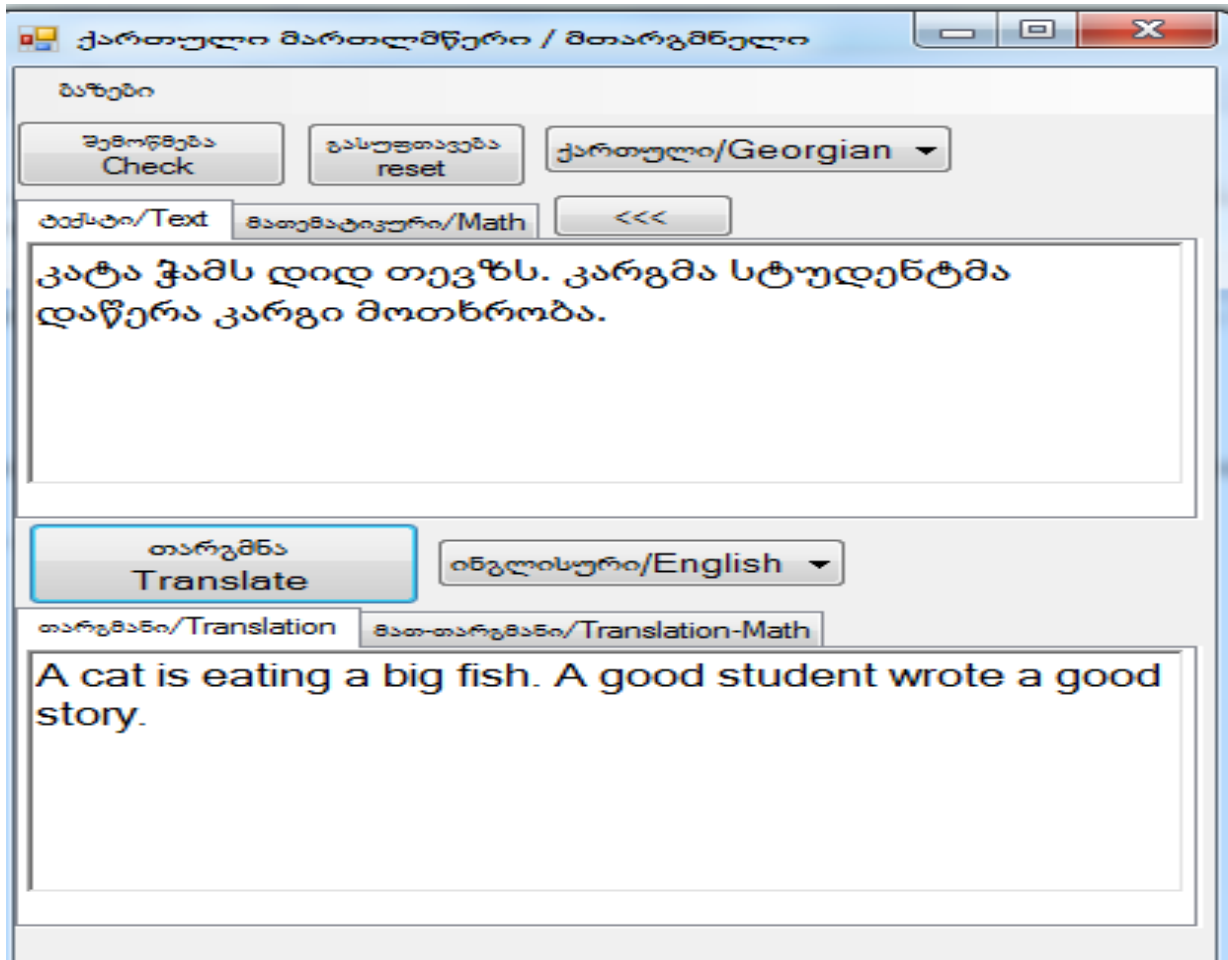
GeoSyntaxAnalyzer/Translator_2009 სისტემის ამ ნაწილობრივი სტრუქტურული აღწერის შემდეგ მისი ფუნქციონირების უფრო თვალსაჩინოდ წარმოჩენის მიზნით ქვემოთ გთავაზობთ ამ ფუნქციონირების ამსახველ სურათებს (იხ., ქვემოთ, სურ. 3 - 8) მცირე კომენტარებითურთ.



სურ. 3

¹⁵ სისტემა, სინტაქსურად აანალიზებს მხოლოდ ქართულ ტექსტებს. ამასთან, სისტემის C ვერსია, რომელიც ამ სისტემის აქ განხილვადი დელფი ვერსიის გადაწერით შეიქმნა, სინტაქსურ შეცდომებზე ხაზგასმის გარდა იძლევა შემოთავაზებებს შეცდომების შესაძლო გასწორებათა თაობაზე.

სურ. 3-დან ჩანს, რომ სისტემამ გაანალიზების შედეგად ხაზი გაუსვა ზედა ველში ჩაწერილი ტექსტის ორივე წინადადებას, რითაც მომხმარებელს მიაჩნია მათში არსებულ სინტაქსურ შეცდომებზე.

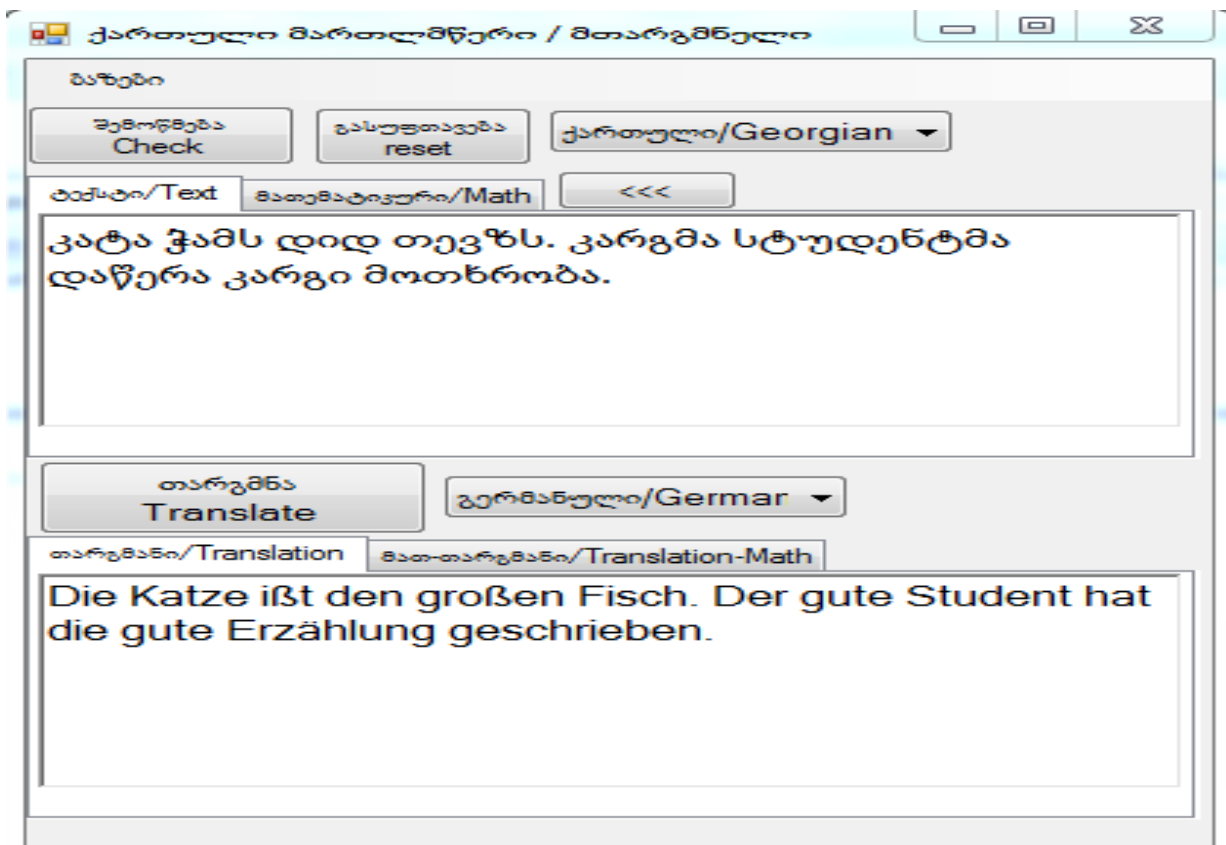


სურ. 4

მომხმარებელმა, მას შემდეგ რაც მან გაასწორა სისტემით მინიშნებული სინტაქსური შეცდომები, გადაწყვიტა, ეს სინტაქსურად უკვე გამართული ქართული ტექსტი ეთარგმნა ინგლისურად (იხ. სურ 4).

ამ მიზნით მან აირჩია ენათა წყვილი ქართული - ინგლისური. ამისთვის მან ზედა ველის ენის გადამრთველი დააყენა **ქართული/Georgain**, ხოლო ქვედა ველის - **ინგლისური/English** მდგომარეობაზე, რის შემდეგაც ქვედა ველის **თარგმნა/Translate** ღილაკის გააქტიურებით მას სისტემამ ქვედა ველში მიაწოდა თარგმანი (იხ. სურ. 4).

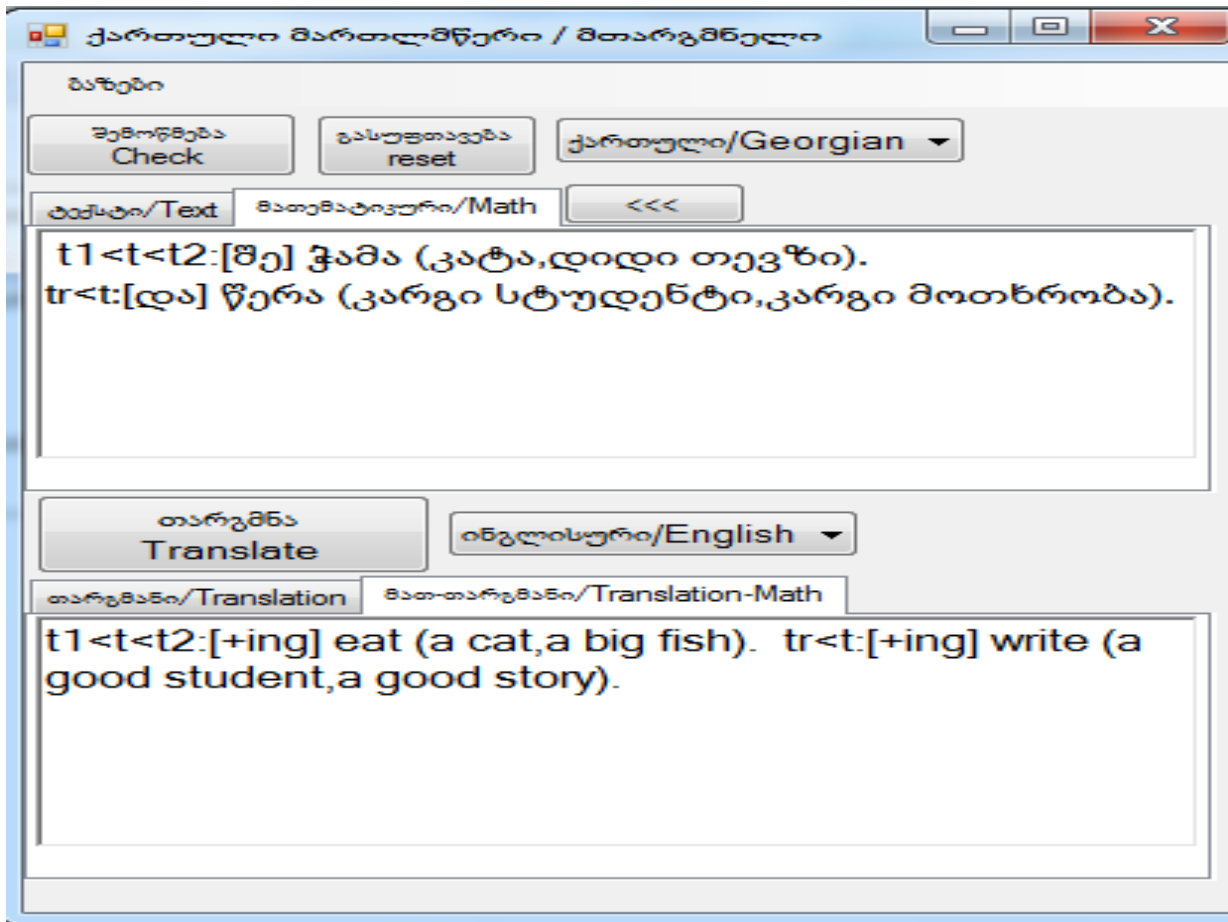
გასაგებია, რომ ანალოგიურად უნდა მოქცეულიყო მომხმარებელი, თუ ის თავიდანვე გადაწყვიტდა გაეკეთებინა თავდაპირველი ქართული ტექსტის გერმანული თარგმანი. ამასთან, თუ მომხმარებელმა უკვე ათარგმნინა სისტემას ტექსტი ინგლისურად, ანუ, თუ სისტემა უკვე იმყოფება სურ. 4-ით აღწერილ მდგომარეობაში, იმავე ქართული ტექსტის გერმანული თარგმანის მისაღებად საკმარისია ქვედა ველის ენის გადამრთველი გადავიყვანოთ **გერმანული/German** პოზიციაზე და გავააქტიუროთ ღილაკი **თარგმნა/Translate** (იხ. სურ. 5).



სურ. 5

ქვემოთ, სურ. 6-ზე, აღწერილია მდგომარეობა, რომელიც მიიღება სურ. 4-ით აღწერილი მდგომარეობიდან ზედა და ქვედა ველების მათემატიკური თარგმანების მომცემი ღილაკების გააქტიურებით. სურ. 6-ით გამოტანილი სისტემის გარეფორმის ზედა ველში წარმოდგენილია თავდაპირველი ქართული ტექსტის მათემატიკური თარგმანი, ხოლო ქვედა ველში წარმოდგენილია ამ ქართული მათემატიკური ტექსტის სიტყვა-სიტყვითი თარგმნის გზით მიღებული ინგლისური მათემატიკური ტექსტი, რაც, თავის მხრივ, სისტემისათვის არის თავდაპირველი ქართული ტექსტის ინგლისური თარგმანის საგენერაციო (წარმოქმნელი) მოცემულობა.

ამგვარად, სისტემაში მათემატიკური ენა და მათემატიკური თარგმანი, ერთი მხრივ, საშუამავლო როლს თამაშობს თარგმნითი ურთიერთკავშირების დამყარებაში, მეორე მხრივ კი, ამ შუალედური მათემატიკური თარგმანის მეშვეობით, სხვადასხვა ლოგიკური და გრამატიკული სტრუქტურის მქონე ენებს შორის სიტყვა-სიტყვითი თარგმნის შეუძლებლობით წარმოქმნილი სირთულეები მოიხსნა მათემატიკურ ენაში პრიმიტიული - სიტყვა-სიტყვითი თარგმანის გაკეთების სიმარტივით, რაც აქ წარმოდგენილი მთარგმნელი სისტემის მეტად მნიშვნელოვანი მახასიათებელია (უფრო დეტალურად იხ. [25], [21], [20], [9], [4]).

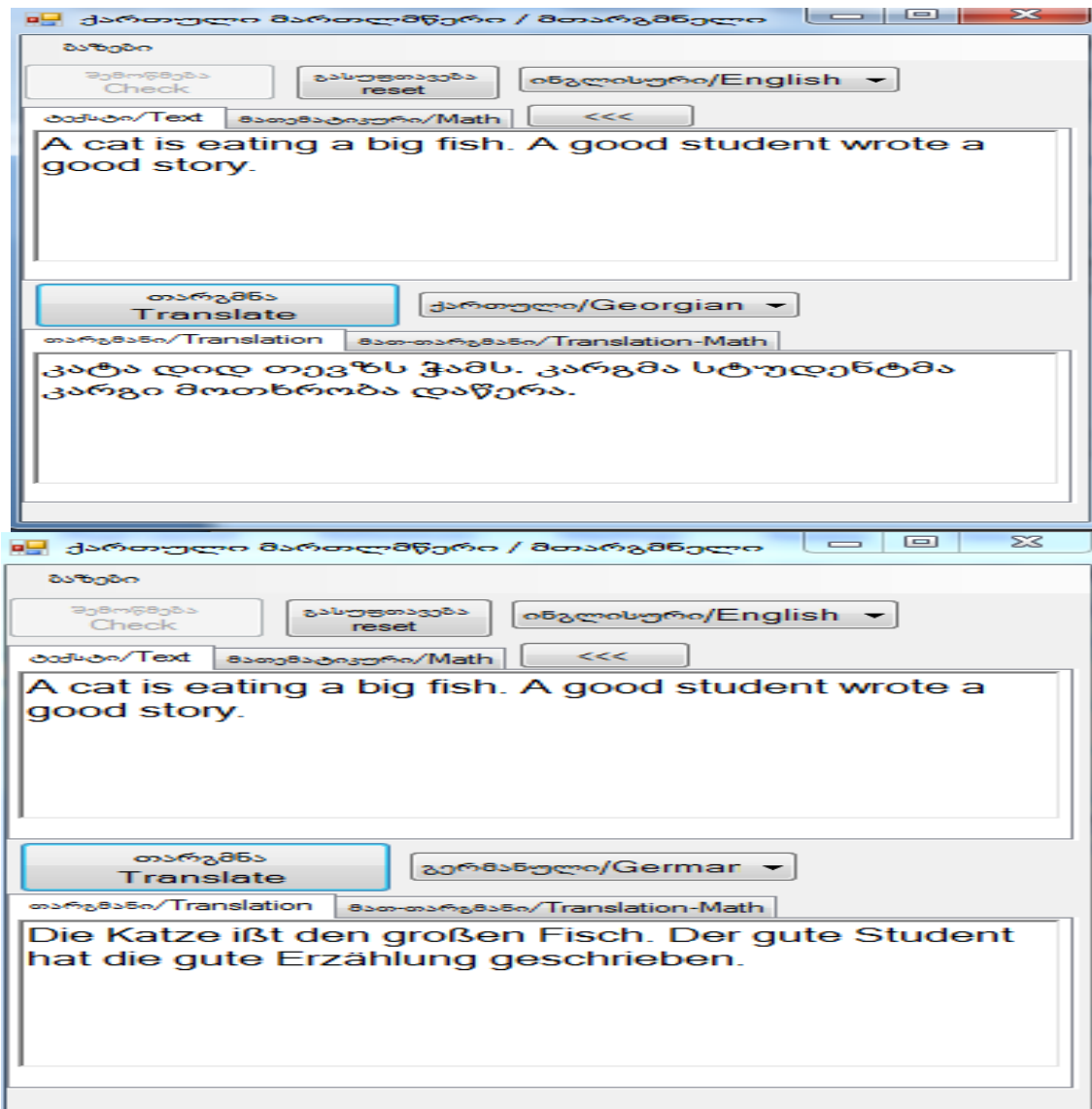


სურ. 6

ქვემოთ, სურ. 7-ზე და 8-ზე, აღწერილია მდგომარეობა, როცა მომხმარებელი სისტემას იყენებს ინგლისურიდან ქართულ და გერმანულ ენებზე მთარგმნელ სისტემად.¹⁶ ასეთ შემთხვევაში, მომხმარებელმა ზედა ველის ენის გადამრთველი უნდა დააყენოს **ინგლისური/English** მდგომარეობაზე, ქვედა ველის ენის გადამრთველი კი - **ქართული/Georgain**, შესაბამისად **გერმანული/German**, მდგომარეობაზე, რის შემდეგაც ქვედა ველის **თარგმნა/Translate** ღილაკის გააქტიურებით ქვედა ველში სისტემა მომხმარებელს მიაწვდის სათარგმნი ინგლისური ტექსტის ქართულ, შესაბამისად გერმანულ, თარგმანს. ზუსტად ანალოგიურად მუშაობს სისტემა იმ შემთხვევაში, როცა მომხმარებელს სჭირდება იგი გერმანულიდან ქართულ, შესაბამისად, ინგლისურ ენაზე მთარგმნელ სისტემად.¹⁷

¹⁶ სისტემის უნარი თარგმნით მიმართებაში ჩაეყენებინა ერთმანეთთან ინგლისური და გერმანული ენები იყო არა ჩვენი მიზანი, არამედ ის „გვერდითი“ ანუ დამატებითი ეფექტი, რაც თარგმნის ამოცანასთან მიმართებაში კ.ფხაკაძის მიერ წინასწარი სახით 2004-2006 წლებში შემუშავებულმა მიდგომებმა გამოიღო (იხ. [30], [25]).

¹⁷ მიუხედავად იმისა, რომ მანქანური თარგმნის საკითხი სცდება საანგარიშო თემატიკას, ქართული ენის ტექნოლოგიების დღევანდელი არასახარბიელო სურათის უფრო სრული წარმოჩენის მიზნით, საჭიროა ხაზი გაესვას იმასაც, რომ აქ მიმოხილული მთარგმნელის სისტემა არის გრამატიკულ, ლოგიკურ და სემანტიკურ ანალიზზე დამყარებულ პირველი ქართული ექსპერიმენტული მთარგმნელი სისტემა. თუმცა, აქ არ შეიძლება არ აღინიშნოს <http://translate.any.ge/> და <http://translate.ge/> მისამართებზე განთავსებული და გუგლზე მიერთებული არასემანტიკური მთარგმნელი სისტემები, როგორც ამ მიმართულებით გადადგმულ პირველ მნიშვნელოვან ქართულ ნაბიჯად.



სურ. 7- 8

3.1.2. ModusPonens_2005 სისტემის აღწერაში, რომელიც თავიდანვე ჩაფიქრებული იყო როგორც ორპირობიანი ლოგიკური ამოცანების ამომხსნელი და შემმოწმებელი ექსპერიმენტული სისტემა, დაგვეხმარება სისტემის გარე ფორმა და მისი ფუნქციონირების ამსახველი სურათები (იხ., ქვემოთ, სურ. 9-13).

ModusPonens_2005 სისტემის სამომხმარებლო გარე ფორმაზე (სურ. 9):

1. 2 პირობიანი ამოცანის ჩასაწერად **CawereT winadadebebi** ფრაზით გამოყოფილია 1 და 2 ციფრებით მონიშნული ორი სტრიქონი;
2. მომხმარებლისეული დასკვნის ანუ ამოცანის მომხმარებლისეული ამოხსნის ჩასაწერად **CawereT daskvna** ფრაზით გამოყოფილია ერთი სტრიქონი;
3. ყვითელი ფერის სტრიქონის ქვედა სტრიქონით სისტემა ამყარებს დიალოგს მომხმარებელთან;
4. ამოცანის ამოხსნის/ამონახსნის შემოწმების მიზნით **Cveni daskvna** ფრაზით გამოყოფილია ერთი სტრიქონი;

ახლა დავუშვათ, რომ მომხმარებელმა ზემოაღწერილი ინსტრუქციის შესაბამისად შეავსო სისტემის სამომხმარებლო გარე ფორმა, ისე როგორც ეს ნაჩვენებია სურ. 9-ზე. ამით მან სისტემას შესამოწმებლად მისცა მის მირ უკვე ამოხსნილი ორპირობიანი ლოგიკური ამოცანა.

| ნაწილები | წინადადებები | | | |
|---|--------------|-----------|-----------|-------|
| 1 | ყველა | სტუდენტი | კითხულობს | წიგნს |
| 2 | გოგო | არის | სტუდენტი | |
| ნაწილები | დასკვნა | | | |
| | გოგო | კითხულობს | წიგნს | |
| | | | | |
| პირველ წინადადებაში სინტაქსური შეცდომაა | | | | |

სურ. 9

ModusPonens_2005 სისტემამ გააანალიზა რა შემოსული დავალება, სადიალოგო ხაზით მომხმარებელს მიაწოდა შეტყობინება ამოსახსნელი ამოცანის პირველ წინადადებაში დაშვებულ სინტაქსურ შეცდომაზე (იხ. სურ. 9).

მომხმარებელმა გაითვალისწინა რა სისტემის მინიშნება, გაასწორა პირველ წინადადებაში დაშვებული სინტაქსური შეცდომა, რის შემდეგაც სისტემამ ამ ამოცანასთან დაკავშირებით დაადასტურა მომხმარებლის დასკვნის მართებულობა (იხ. სურ. 10).

| ნაწილები | წინადადებები | | | |
|----------------|--------------|-----------|-----------|-------|
| 1 | ყველა | სტუდენტი | კითხულობს | წიგნს |
| 2 | გოგო | არის | სტუდენტი | |
| ნაწილები | დასკვნა | | | |
| | გოგო | კითხულობს | წიგნს | |
| | | | | |
| დასკვნა სწორია | | | | |
| ნაწილი | დასკვნა | | | |
| ეს | გოგო | კითხულობს | წიგნს | |

სურ. 10

სურ. 11-ზე და 12-ზე ნაჩვენებია ორი ისეთი ერთმანეთისგან განსხვავებული ვითარება, როცა სისტემა ატყობინებს მომხმარებელს დასკვნის არამართებულობას და, ამავდროულად, აწვდის მას სწორ დასკვნას. ამგვარად, ამ ორივე შემთხვევაში, სისტემა მომხმარებელს უხსნის მის მიერ არასწორად ამოხსნილ ლოგიკურ ამოცანებს.

| ნაწილები | წინადადებები | | | |
|-------------------|--------------|-----------|-----------|-------|
| 1 | ყველა | სტუდენტი | კითხულობს | წიგნს |
| 2 | გოგო | არის | სტუდენტი | |
| ნაწილები | დასკვნა | | | |
| | ბიჭი | კითხულობს | წიგნს | |
| | | | | |
| დასკვნა არასწორია | | | | |
| ნაწილი | დასკვნა | | | |
| ეს | გოგო | კითხულობს | წიგნს | |

სურ. 11

| | | | | |
|---------|-------------------|-----------|-----------|-------|
| ჩაწერეთ | წინადადებები | | | |
| 1 | ყველა | სტუდენტი | კითხულობს | წიგნს |
| 2 | ბიჭი | არის | სტუდენტი | |
| ჩაწერეთ | დასკვნა | | | |
| | ბიჭი | კითხულობს | სახლს | |
| | | | | |
| | დასკვნა არასწორია | | | |
| ჩვენც | დასკვნა | | | |
| ეს | ბიჭი | კითხულობს | წიგნს | |

სურ. 12

სურ. 13-ზე სისტემა ატყობინებს მომხმარებელს იმის თაობაზე, რომ მისი ამოცანის წინაპირობით რაიმე სახის კერძო დასკვნის გაკეთება შეუძლებელია.

| | | | | |
|---------|-------------------------------|----------|-----------|-------|
| ჩაწერეთ | წინადადებები | | | |
| 1 | ზოგიერთი | სტუდენტი | კითხულობს | წიგნს |
| 2 | გოგო | არის | სტუდენტი | |
| ჩაწერეთ | დასკვნა | | | |
| | | | | |
| | დასკვნის გაკეთება შეუძლებელია | | | |

სურ. 13

ზემოაღწერილი ModusPonens_2005 სისტემა, რომელიც 2005 წელს შემუშავდა, არის პირველი ექსპერიმენტი, სადაც ქართული ტექსტობრივი მასალის სინტაქსური და სემანტიკური (ლოგიკური) ანალიზი გაკეთდა ამ ტექსტის მათემატიკურ ენაზე წინასწარი თარგმნისა და უკვე მათემატიკურ ენაში მისი დამუშავების გზით.

3.1.3. GeoNetSyntaxAnalyzer_2011 სისტემის აღწერაში დაგვეხმარება ტექსტში ჩაბეჭდილი სურათები. კერძოდ, GeoNetSyntaxAnalyzer_2011 სისტემის საწყის გვერდს შემდეგი სახე აქვს (იხ., ქვემოთ, სურ. 14).



სურ. 14

სურათზე გამოტანილი ფორმის მარჯვენა ველში იწერება გასაანალიზებელი ქართული ტექსტი, რომელიც აგებული უნდა იყოს:

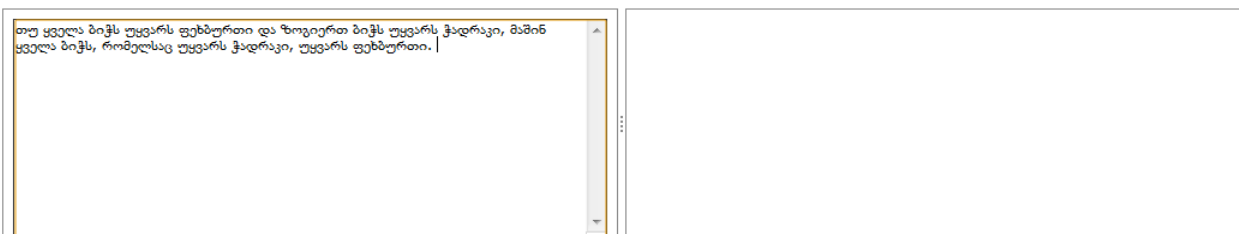
1. ქართული ენის ლოგიკური გრამატიკის მარტივი წინადადებებით;
2. ქართულ ენის რთული წინადადებების ამგები ძირეული ლოგიკური ოპერატორებით. ესენია:
 1. და მარცხნივ და მარჯვნივ 1-ადგილიანი კავშირი;
 2. ან ,ან ტიპის კავშირთა სისტემა, რომლებიც აგებულია ან მარჯვნივ 1-ადგილიანი საკავშირო ელემენტისგან;
 3. თუ ,მაშინ ორადგილიანი შედგენილი ლოგიკური კავშირი, რომელიც აგებულია მარჯვნივ 1-ადგილიანი თუ და მაშინ საკავშირო ელემენტისაგან;
3. ქართული ენის ზმნური სახელური ფრაზების ამგები რომელიც ტიპის მარჯვნივ და მარცხნივ 1-ადგილიანი ლოგიკო-სპეციალური ოპერატორებით.

სურ. 15-ზე ნაჩვენებია სისტემის ის მდგომარეობა, როცა სისტემის მარცხენა ფანჯარაში ჩაწერილია სწორედ ზემოაღნიშნული ენობრივი საშუალებებით აგებული ქართული ტექსტი.¹⁸

სინტაქსური ანალიზის დილაკის გააქტიურებით სისტემა აანალიზებს მარცხენა ველში განთავსებულ ტექსტს და ანალიზის შედეგებს შეფერადებული გადაწერით წარმოგვიდგენს მის მარჯვენა ველში (იხ., ქვემოთ, სურ. 16).

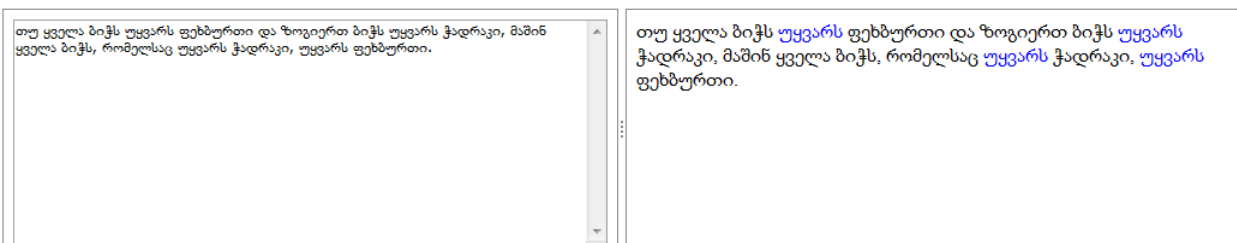
[ბაზები](#)

სინტაქსური ანალიზი



[ბაზები](#)

სინტაქსური ანალიზი



სურ. 15-16

¹⁸ ტექსტი ლიტერატურულია. - ვგულისხმობთ იმას, რომ აქ და ქვემოთაც ტექსტების ამგებ წყაროდ აღებულია შეფასებისა და გამოცდების ეროვნული ცენტრის ზოგადი უნარების ერთიანი ეროვნული გამოცდების სხვადასხვა წელს გამოცემული დამხმარე მასალათა კრებულები, სადაც სტანდარტიზებული სახით მრავლადაა მოცემული საკმაოდ რთული ლოგიკური შინაარსის მქონე ქართული წინადადებები.

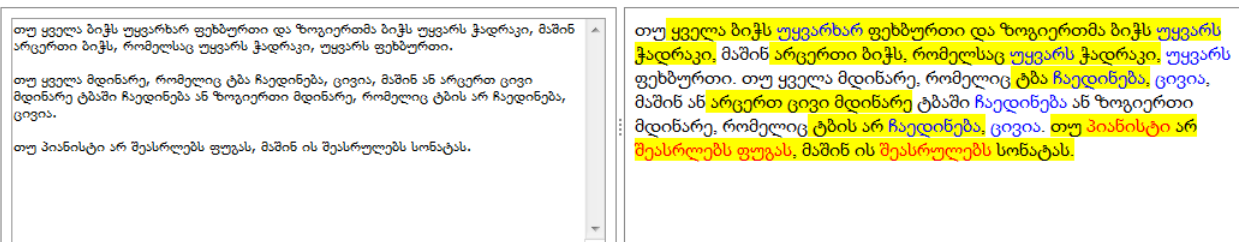
ამასთან, იმ შემთხვევაშიც, თუ გასაანალიზებელ ტექსტში სისტემის მიერ ვერ იქნა შემჩნეული ვერცერთი შეცდომა, მარჯვენა ველში გადაწერილ ტექსტში ლურჯად ფერადდება ზმნები (სურ. 16). ეს ითვალისწინებს ამ სისტემაზე ჯგუფშივე შემუშავებული სინთეზატორების მიზმის მიზნებს, და წასაკითხი ტექსტის სინთეზირებისას კითხვის ხარისხის ამაღლებას.¹⁹

რაც შეეხება იმ შემთხვევას, როცა გასაანალიზებელი ტექსტის ანალიზის შედეგად სისტემის მიერ დაძებნილ იქნა სინტაქსური შეცდომა ან სისტემისთვის უცნობი სიტყვა, რაც ამ სისტემას ორთოგრაფის ფუნქციასაც ანიჭებს, მარჯვენა ველში გადაწერილი ტექსტის ის ადგილები, სადაც სისტემამ ეს სინტაქსური ან შესაძლო ორთოგრაფიული შეცდომა დაძებნა, ფერადდება ყვითლად. სისტემა მისთვის უცნობ სიტყვებს აფერადებს წითლად.²⁰ შეფერადებულ ადგილებზე კურსორის (თაგვისარას) მიტანით, სისტემა მომხმარებელს აწვდის ტექსტობრივ შეტყობინებას ტექსტის ამ შეფერადებული ნაწილის შეფერადების მიზეზის თაობაზე. ამასთან, როგორც უკვე აღვნიშნეთ, ეს შეტყობინება არის ან შეტყობინება ტექსტის შეფერადებულ ნაწილში უცხო სიტყვის არსებობაზე, ანდა შეტყობინება იმ სინტაქსურ შეცდომაზე, რაც შეფერადებულ ნაწილში სისტემამ შეამჩნია.

აქ აღნიშნულის გათვალისწინებით, ქვემოთ, სისტემის სხვადასხვა სახის ფუნქციონირების ამსახველი სურათებითა და მათზე თანრთული კომენტარებით (იხ., ქვემოთ, სურ. 17 - 61), მოკლედ და თითქმის ამომწურავად აღვწერთ სისტემას ფუნქციონალური და სტრუქტურული თვალსაზრისით.

ბაზები

სინტაქსური ანალიზი



სურ. 17

სურ. 17-დან ჩანს, რომ სისტემამ, გააანალიზა სამი რთული წინადადებით აგებული ტექსტი. ამასთან:

¹⁹ ანუ, ეს ჩაფიქრებულია მკითხველი სისტემის არასემანტიკური (ე.ი. გაუაზრებელი) კითხვიდან სემანტიკურ (ანუ გააზრებულ) კითხვაზე თანდათანობითი გადაყვანის მიზნით, რაშიც, ცხადია, მარტივ წინადადებებში მათი ამგები ზმნის გამოყოფა, ერთ-ერთი აუცილებელი ფაქტორია.

²⁰ ამით სისტემა, თავის თავზე ორთოგრაფის ფუნქციასაც იღებს. - ანუ, მომხმარებელმა უკვე იცის, რომ ის სიტყვა, რომელიც სისტემამ წითლად არ შეაფერადა, ორთოგრაფიულად გამართულია.

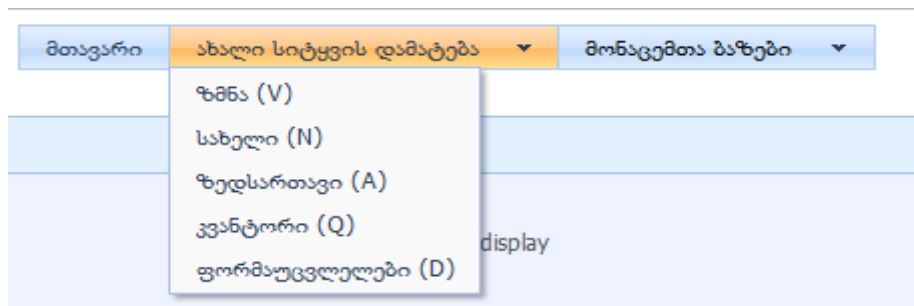
ამასთან, უცხო სიტყვის გამოცნობის ეს უნარი, სისტემას, მას შემდეგ, რაც იგი ქართულ ენაში ლექსიკურად გასრულდება და ამუშავდება, საშუალებას მისცემს აღრიცხოსა და დაიმახსოვროს მასში და, ამდენად, ქართულ ენაში შემოსული ახალი სიტყვის შემოსვლის ზუსტი თარიღი. ეს, ცხადია, მომავალ თაობებს გაუმარტივებს ქართული სიტყვის ისტორიოგრაფიულ კვლევას. - ანუ, ფაქტიურად, ასეთი ქართული სისტემის ამუშავების შემდეგ აქ არც არაფერი აღარ დარჩება საკვლევია, გარდა თუ არა იმ სიტყვების ისტორიოგრაფიისა, რაც ამ აქ ნახსენები ქართული სისტემის ამუშავებამდე უკვე იყო ქართულ ენობრივ-კულტურულ გარემოში.

1. მან დაიჭირა სინტაქსური შეცდომა პირველი წინადადების ორივე ქვეწინადადებაში. ესენია: ‘ყველა ბიჭს უყვარხარ ფეხბურთი’ და ‘ზოგიერთმა ბიჭს უყვარს ჭადრაკი’;

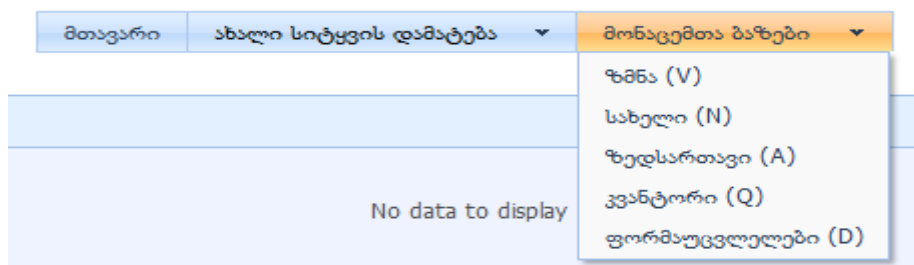
2. მან დაიჭირა სინტაქსური შეცდომები მეორე წინადადების სხვადასხვა ქვეწინადადების სხვადასხვა სახელურ ფრაზებში. ესენია: არცერთი ბიჭს, რომელსაც უყვარს ჭადრაკი , ყველა მდინარე, რომელიც ტბა ჩაედინება, არცერთ ცივი მდინარე და ‘ზოგიერთი მდინარე, რომელიც ტბის არ ჩაედინება’;

3. მესამე წინადადებაში მან ოთხ ადგილზე დაიჭირა ოთხი უცხო სიტყვა და წითლად შეაფერადა. ესენია: ‘პიანისტი’, ‘შეასრულებს’, ‘ფუფას’, ‘შეასრულებს’

იმისათვის, რომ საანალიზო ტექსტმა სრულად გაიაროს მართლმწერი სისტემა, ანუ იმისათვის რომ სისტემამ ტექსტის მესამე წინადადებაც გააანალიზოს, მომხმარებელმა სისტემას უნდა ასწავლოს ამ წინადადებაში შემავალი და მისთვის ჯერ კიდევ უცნობი სიტყვები ‘პიანისტი’, ‘შეასრულებს’, ‘ფუფას’, ‘შეასრულებს’. ამას სისტემა ანხორცილებს სისტემაში ახალ საბაზო მონაცემად ახალი სიტყვის დამატების ავტომატური მექანიზმებით (იხ, ქვემოთ, სურ. 18 და 19)



სურ. 18



სურ. 19

ამასთან, ამჯერად, სისტემაში ინტეგრირებულია 5 ბაზა. ესენია: **V-ბაზა** ანუ **ზმნების ბაზა**, **N-ბაზა** ანუ **სახელების ბაზა**, **A-ბაზა** ანუ **ზედსართავების ბაზა**, **Q-ბაზა** ანუ **განმაზოგადებლების (კვანტორების) ბაზა** და **D-ბაზა** ანუ **ფორმაუცვლელი სიტყვების ბაზა** (იხ., ზემოთ, სურ. 19).

შესაბამისად, სისტემაში, თითოეულ ამ ბაზაზე (გარდა ფორმაუცვლელი სიტყვების ბაზისა) მიდგმულია საბაზო მონაცემების ავტომატურად მაწარმოებელი (მაგენერირებელი) მექანიზმი. კერძოდ, სისტემაში უკვე ინტეგრირებულია **A_გენერატორი** (სურ. 20), **Q_გენერატორი** (სურ. 21), **N_გენერატორი** (სურ. 22), **V_გენერატორი** (სურ. 23). - ქვემოთ, მინიშნებული სურათების დახმარებით, მოკლედ აღვწერთ მათ ფუნქციონალური და სტრუქტურული თვალსაზრისებით:

1. A_გენერატორში (Q_გენერატორში), ანუ ზედსართავი (განმაზოგადებელი) სახელების საბაზო ბრუნვითი ფორმების მაწარმოებელ სისტემაში (იხ., ქვემოთ, სურ. 20 და სურ. 21) შეგვყავს შესაბამისი ტიპის საბრუნებელი სიტყვა და წითელი ხაზის გასწვრივ მდებარე ღილაკის გააქტიურებით ვანხორციელებთ მის ბრუნებას. **ბაზაში დამატება** ღილაკის გააქტიურებით ნაბრუნები ფორმები ემატება ბაზას ამ ბაზის შესაბამისი ტიპის მათემატიკურად აღწერილ ახალ მონაცემებად, რის შემდეგაც სისტემა აფართოებს ამ ახალ სიტყვებზე თავის მოქმედებას. ეს სისტემას საშუალებას აძლევს ძველ ტექსტებთან ერთად გააანალიზოს ამ ახლად ჩადებული ანუ ნასწავლი სიტყვებით შედგენილი ახალი ტექსტებიც.²¹

სურ. 20
A_გენერატორი

სურ. 21
Q_გენერატორი

2. N_გენერატორი, ანუ არსებითი სახელების საბაზო ბრუნვითი ფორმების მაწარმოებელი სისტემა (სურ. 22-ზე წარმოდგენილია ამ სისტემის მხოლოდ ერთი - ძირითადი ქვესისტემის გარეფორმა), რეალიზებას უკეთებს თითქმის ნებისმიერი არსებითი სახელის ბრუნებას ქართული ენის ლოგიკური გრამატიკით შემუშავებული სისტემის მიხედვით.²²

²¹ სისტემები აგებულია ქართული ენის ლოგიკური გრამატიკით შემუშავებული სქემების მიხედვით (იხ. [30]).

²² იხ. [30] - კ.ფხაკაძის 2005 წელს შემუშავებულ ლოგიკურ ბრუნებასა და ტ.ფუტკარაძის 1997 წელს შემოთავაზებულ ბრუნვით სისტემებს შორის (იხ. [31]) მეტად მჭიდრო ფორმობრივი მსგავსებაა. ამასთან, კ.ფხაკაძის სისტემა, რომელიც ქართული და მათემატიკური ენების ბუნებრივი შინაარსული ურთიერთმიმართებების კვლევის მიზნითაა აგებული, მისი ამგები მათემატიკური - ლოგიკური პრინციპით უფრო ახლოსაა რუსულ ბრუნებასთან დაკავშირებულ ა.ნ.კოლმოგოროვისეული ხედვის მათემატიკურ პრინციპთან (იხ. [52]), თუმცა მისგანაც არსებითად განსხვავდება, ვიდრე

სახელი (N)

| | | | |
|---------------------------------------|----------------------|--|----------------------|
| <input type="button" value="აბრუნე"/> | | ფუძე(1) <input type="text"/> | |
| | | ფუძე(2) <input type="text"/> | |
| N-ი | <input type="text"/> | N-ს | <input type="text"/> |
| N-ის | <input type="text"/> | N-ში | <input type="text"/> |
| N-ისთვის | <input type="text"/> | N-ზე | <input type="text"/> |
| N-ისკენ | <input type="text"/> | N-ად | <input type="text"/> |
| N-ისგან | <input type="text"/> | N-მდე | <input type="text"/> |
| N-იდან | <input type="text"/> | N-თან | <input type="text"/> |
| N-ით | <input type="text"/> | N-მა | <input type="text"/> |
| <input type="checkbox"/> სულიერი | | <input type="button" value="ბაზაში დამატება"/> | |

სურ. 22
N_გენერატორი

ამის შემდეგ, **ბაზაში დამატება** ღილაკის გააქტიურებით ნაბრუნები ფორმები ემატება ბაზას ამ ბაზის შესაბამისი ტიპის მათემატიკურად აღწერილ ახალ მონაცემებად, რის შემდეგაც სისტემა აფართოებს ამ ახალ სიტყვებზე თავის მოქმედებას. ეს სისტემას საშუალებას აძლევს ძველ ტექსტებთან ერთად გააანალიზოს ამ ახლად ჩადებული ანუ ნასწავლი სიტყვებით შედგენილი ახალი ტექსტებიც.

2. **V_გენერატორი, ანუ ზმნების საბაზო უღლებითი ფორმების მაწარმოებელი სისტემა** (იხ., ქვემოთ, სურ. 23, რომელზეც წარმოდგენილია ამ სისტემის გარე ფორმა) ქართული ენის ლოგიკური გრამატიკით შემუშავებული სქემების მიხედვით რეალიზებას უკეთებს თხრობითი კილოს მხოლობითი რიცხვის თითქმის ნებისმიერი ზმნის უღლებასა და უღლების პარალელურ მათემატიკურ აღწერას 4 დროში.²³ ამასთან,

ტ.ფუტკარაძისეული მიდგომის განმსაზღვრელ პრინციპთან. - მიუხედავად ამისა, ჯგუფში მიმდინარეობს მუშაობა ამ ორი - ერთმანეთთან ფორმოზირებად მეტად ახლო მიდგომი სისტემების შეჯერების მიმართულებით. - ამ თვალსაზრისით, გასათვალისწინებელია ლ.აბზიანიძის ხედვაც, რომელიც კ.ფხაკაძის 14 ფორმიან ბრუნებას უმატებს კიდევ ერთ - კუთვნილებითი მიმართების გამომხატველ ფორმას (იხ. [3]), რომელიც [30]-ში ცალკე მორფო-სინტაქსურ ტიპად არის გამოყოფილი. - გასაგებია, რომ არსებითი სახელის ბრუნებაზე გამოკვეთილი ეს ახალი ხედვები და მათზე დამყარებული ლინგვისტური მიდგომები არსებითად განსხვავდება დღეს გაბატონებული კლასიკური ხედვებისაგან (იხ. [51], [47], [43]) და მათზე დამყარებული მიდგომებისაგან ([48], [46], [39], [37], [15]).

²³ (იხ. [32], [29]). - აქ უნდა აღინიშნოს, ქართული ენის ლოგიკური გრამატიკის ფარგლებში 2004-2005 წლებში შემუშავებული კ.ფხაკაძის ლოგიკური უღლებისა და ქართული ზმნის უღლებაზე ტ.ფუტკარაძის ხედვის თვალშისაცემი მსგავსება (იხ. [42], აგრეთვე, ტ.ფუტკარაძის ნაშრომი „ქართული ზმნის ლოგიკური პარადიგმის საკითხისათვის“ (აკაკი წერეთლის სახელმწიფო უნივერსიტეტის ჰუმანიტარულ მეცნიერებათა ფაკულტეტის შრომები, ენათმეცნიერული სერია, ტ.VIII (I), ქუთაისი, 2006

ბაზაში დამატება ღილაკის გააქტიურებით სისტემის მიერ უკვე ნაუღლები ფორმები ემატება ბაზას ამ ბაზის შესაბამისი ტიპის მათემატიკურად აღწერილ ახალ მონაცემებად, რის შემდეგაც სისტემა აფართოებს ამ ახალ სიტყვებზე თავის მოქმედებას. ეს სისტემას საშუალებას აძლევს ძველ ტექსტებთან ერთად გააანალიზოს ამ ახლად ჩადებული ანუ ნასწავლი სიტყვებით შედგენილი ახალი ტექსტებიც.

სურ. 23
V_გენერატორი

V_გენერატორი, რომელიც სისტემის მეტად მნიშვნელოვანი ქვესისტემაა, თავის თავში მოიცავს ზმნის გენერაციის ანუ წარმოქმნის წესის ამგებ ქვესისტემას. კერძოდ, V_გენერატორის გარე ფორმაზე არსებული **ზმნის გენერაციის წესების ცხრილის** ღილაკის გააქტიურებით (იხ., ზემოთ, სურ. 23) მომხმარებელზე გამოდის **ზმნის უღლების წესების მაწარმოებელი სისტემა** (იხ., ქვემოთ, სურ. 24).

სურ. 24
ზმნის უღლების წესების მაწარმოებელი სისტემა

ეს სისტემა, რომელიც ლოგიკური გრამატიკით შემუშავებული სქემების მიხედვით არის აგებული (იხ. [32], [29]), იძლევა სისტემაზე ახლი უღლების წესის

წელი, გვ.88-96) და ორივე ამ სისტემის ასევე თვალშისაცემი განსხვავებულობა ქართული ზმნის უღლების სისტემაზე წინა წლებში არსებული ხედებისაგან. - ამ შემთხვევაშიც, ჯგუფში უკვე იყო საუბრები ამ ორი მსგავსი, მაგრამ რიგ საკითხებში ერთმანეთისგან პრინციპულად განსხვავებული ხედვის შეჯერების საჭიროებაზე.

მიერთების საშუალებას, რაც სისტემის ავტომატური გაფართოების შესაძლებლობებს თითქმის ზღვრულად აფართოებს.

ამგვარად, ჩვენ უკვე მოკლედ მიმოვიხილეთ GeoNetSyntaxAnalyzer_2011 სისტემის რთული მრავალქვესისტემიანი და მრავალბაზიანი სტრუქტურა. ახლა დავუბრუნდეთ სისტემის ფუნქციონალურ მხარეს და უფრო დეტალურად ვნახოთ როგორ გაართმევს სისტემა თავს სურ. 17-ზე გამოტანილი ტექსტის სრული გაანალიზების ამოცანას.

როგორც უკვე აღვნიშნეთ, ამისათვის მომხმარებელმა სისტემას უნდა ასწავლოს სიტყვები ‘პიანისტი’, ‘ფუფას’, ‘შეასრულებს’. (მომხმარებელი ცხადია მიხვდა, რომ ‘შეასრულებს’ არის ‘შეასრულებს’ ზმნის ორთოგრაფიულად დამახინჯებული ფორმა). ამ მიზნით:

1. მომხმარებელი იძახებს N_გენერატორს (არსებით სახელთა მაწარმოებელს) და საბრუნებლად აძლევს მას სიტყვებს ‘პიანისტი’ და ‘ფუფა’ (სურ. 25-26);

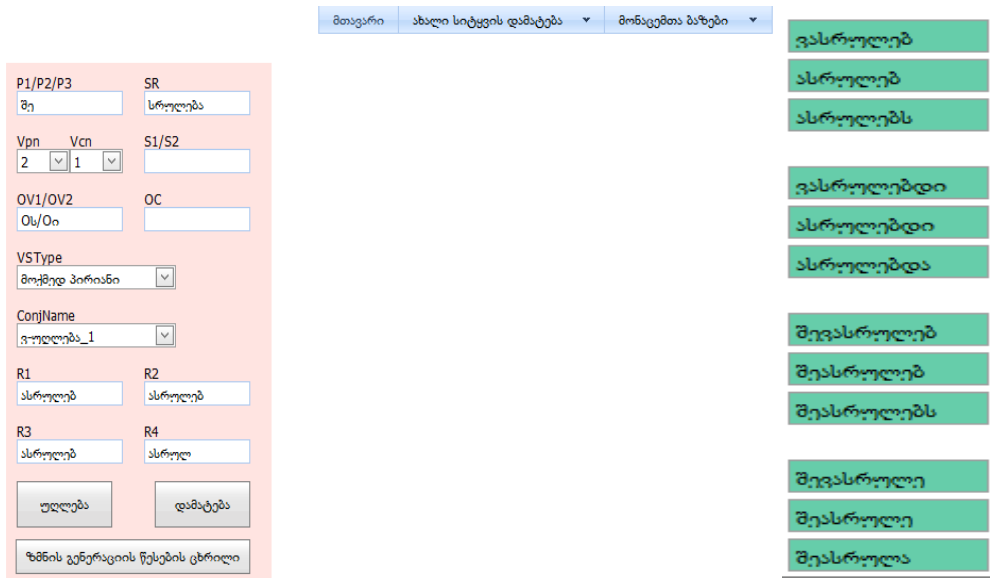
The image shows two side-by-side screenshots of a web application interface for generating N-words. Both screenshots have a title 'სახელი (N)'.
 The left screenshot shows a header with 'აბრუნე' and two input fields: 'ფუძე(1) პიანისტი' and 'ფუძე(2) პიანისტი'. Below is a grid of 12 text boxes for different grammatical forms: N-ი (პიანისტი), N-ს (პიანისტს), N-ის (პიანისტის), N-ში (პიანისტში), N-ისთვის (პიანისტისთვის), N-ზე (პიანისტზე), N-ისკენ (პიანისტისკენ), N-ად (პიანისტად), N-ისგან (პიანისტისგან), N-მდე (პიანისტამდე), N-იდან (პიანისტიდან), N-თან (პიანისტთან), N-ით (პიანისტით), and N-მა (პიანისტმა). At the bottom, there is a checked checkbox 'სულიერი' and a button 'ბაზაში დამატება'.
 The right screenshot shows a similar interface but with different input words: 'ფუძე(1) ფუფა' and 'ფუძე(2) ფუფა'. The corresponding grid of text boxes contains: N-ი (ფუფა), N-ს (ფუფას), N-ის (ფუფის), N-ში (ფუფაში), N-ისთვის (ფუფისთვის), N-ზე (ფუფაზე), N-ისკენ (ფუფისკენ), N-ად (ფუფად), N-ისგან (ფუფისგან), N-მდე (ფუფამდე), N-იდან (ფუფიდან), N-თან (ფუფასთან), N-ით (ფუფით), and N-მა (ფუფამ). At the bottom, the 'სულიერი' checkbox is unchecked and the 'ბაზაში დამატება' button is present.

სურ 25-26

მას შემდეგ, რაც სისტემამ აბრუნა ეს სიტყვები, მომხმარებელი, **ბაზაში დამატება** ღილაკის გააქტიურებით, სისტემას აწვდის ამ სიტყვების სურ. 25-26-ზე გამოტანილ ყველა ფორმას მათივე ტიპობრივი მათემატიკური აღწერებით, რითაც მომხმარებელი ამთავრებს სისტემისათვის ამ ორი ახალი სიტყვის სწავლებას.

2. მომხმარებელი იძახებს V_გენერატორს და სისტემისათვის ‘შეასრულებს’ ზმნის სწავლების მიზნით ავსებს მას ისე როგორც ეს არის სურ. 27-ზე, რის შემდეგაც **უღლება**

ლილაკის გააქტიურებით სისტემა წარმოქმნის 12 ზმნას, რომელთა შორისაა შეასრულებს ზმნაც (სურ. 27).



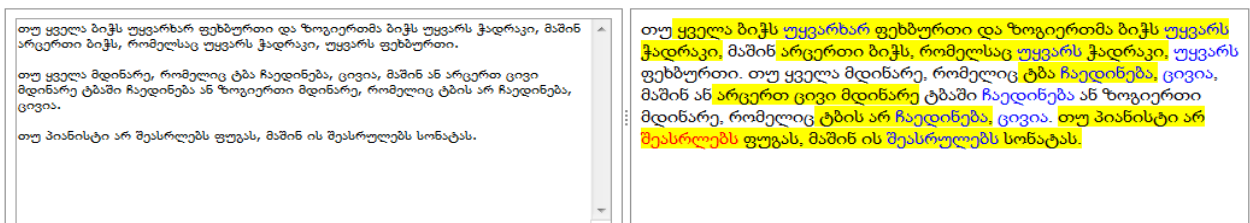
სურ. 27

მას შემდეგ, რაც სისტემამ ვ-ულლება_1 წესით წარმოქმნა სურ 27-ზე გამოტანილი ზმნები და მათი მათემატიკური აღწერები, მომხმარებელი, ბაზაში დამატება ლილაკის გააქტიურებით, სისტემას აწვდის სურ. 27-ზე გამოტანილ ყველა ფორმას მათივე ტიპობრივი მათემატიკური აღწერებით, რითაც მომხმარებელი ამთავრებს სისტემისათვის შეასრულებს სიტყვის სწავლებას.

ახალი სიტყვების სწავლების ამ საკმაოდ მარტივი პროცედურების გავლის შემდეგ მომხმარებელი კვლავ ააქტიურებს სისტემის გარე ფორმის სინტაქსური ანალიზის ლილაკს, რის შემდეგაც სურ 17-ზე არსებული მდგომარეობა იცვლება ქვემოთ სურ 28-ზე გამოტანილი მდგომარეობით.

[ბაზები](#)

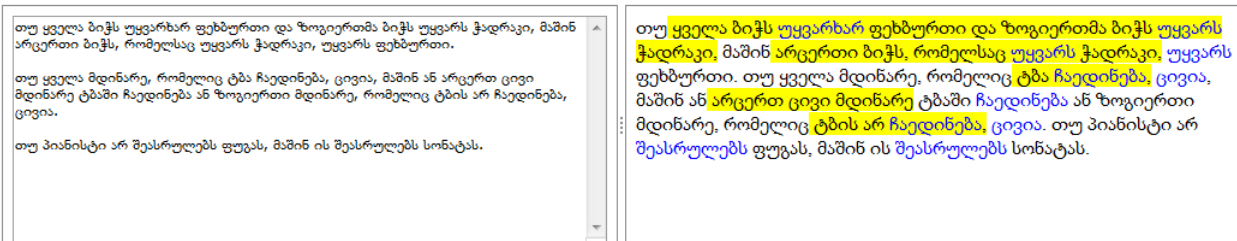
სინტაქსური ანალიზი



სურ. 28

როგორც სურ. 28-დან ჩანს, სისტემამ, ძველ სინტაქსურ შეცდომებთან ერთად მარჯვენა ველში გამოიტანა წითლად შეფერადებული სიტყვის ფორმის ასოთა რიგი შეასრულებს, რაც უკვე შესაძლებელია გასწორდეს ორთოგრაფიულად, რადგან იგი არის სისტემისათვის უკვე ნაცნობი შეასრულებს ზმნის არასწორი ორთოგრაფიული ფორმა.

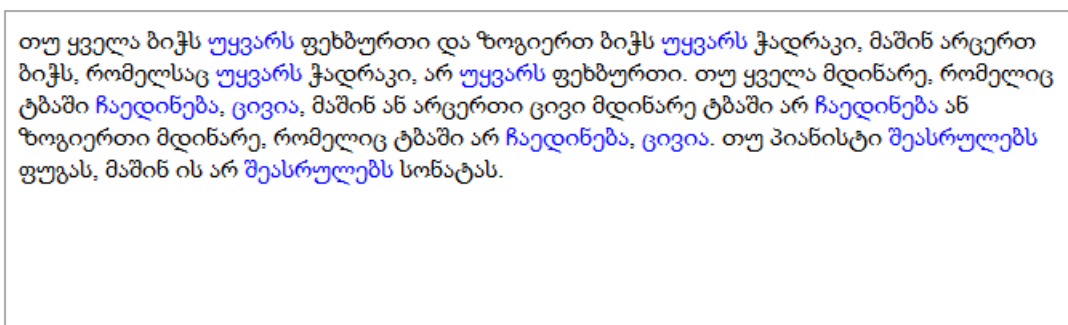
სინტაქსური ანალიზი



სურ. 29

სურ. 29-ზე ასახულია სისტემის ის მდგომარეობა, როცა მომხმარებელმა ტექსტში **შეასრულებს** ფორმა ჩაანაცვლა მისი ორთოგრაფიულად სწორი ვარიანტით, რის შემდეგაც სისტემამ სრულად გააანალიზა მანამდე მისი ცოდნის მიღმა დარჩენილი წინადადება 'თუ პიანისტი არ შეასრულებს ფუგას, მაშინ ის შეასრულებს სონატას'. - ანუ, სისტემამ, სულ რაღაც ხუთწუთიანი პროცედურით, არამარტო ორთოგრაფიულად ისწავლა საანალიზო ტექსტის მესამე წინადადების სამი უცხო სიტყვა, არამედ აითვისა მათი სინტაქსური ფუნქციონირებაც და ეს მანამდე მისთვის სრულიად უცხო წინადადება დაადასტურა როგორც სინტაქსურად გამართული.

ამის შემდეგ მომხმარებელმა გადაწყვიტა გაესწორებინა სისტემის მიერ შემჩნეული სხვა შეცდომებიც, რის შემდგომაც სისტემამ მომხმარებელს დაუდასტურა ტექსტის სრული სინტაქსური გამართულობა (იხ., ქვემოთ, სურ. 30).



სურ. 30

იმ მიზნით, რომ თვალნათელი გავხადოთ სისტემის ფუნქციონირების სხვადასხვა წახნაგი, ქვემოთ, სხვადასხვა გვარად ვარღვევთ სურ. 30-ზე გამოტანილი ტექსტის სინტაქსის. ამასთან, იმის სანახავად, თუ როგორ რეაგირებს სისტემა ამ სხვადასხვა სახის სინტაქსურ დარღვევებზე, ჩვენ დავიხმართ ამ რეაგირებათა ამსახველ სურათებს (იხ., ქვემოთ, სურ. 31-61).

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ არცერთ ბიჭს, რომელსაც უყვარს ჭადრაკი, არ უყვარს ფეხბურთი. თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან არცერთი ცივი მდინარე ტბაში არ ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია თუ პიანისტი შეასრულებს ფუგას, მაშინ ის არ შეასრულებს სონატას.

სურ. 31

სისტემამ შეამჩნია ტექსტის მეორე და მესამე წინადადებებს შორის წერტილის ნაკლებობა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ არცერთ ბიჭს, რომელსაც უყვარს ჭადრაკი, არ უყვარს ფეხბურთი. თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან არცერთი ცივი მდინარე ტბაში არ ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია და თუ პიანისტი შეასრულებს ფუგას, მაშინ ის არ შეასრულებს სონატას.

სურ. 32

მას შემდეგ, რაც გამოტოვებული წერტილის ნაცვლად დაესვით 'და' კავშირი, სისტემამ დაადასტურა ტექსტის სინტაქსის გამართულობა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ არცერთ ბიჭს, რომელსაც უყვარს ჭადრაკს, არ უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან არცერთი ცივი მდინარე ტბაში არ ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია და თუ პიანისტი შეასრულებს დიდი ფუგას, მაშინ ის არ შეასრულებს ზოგიერთი კატარა სონატას.

სურ. 33

აქ გასაანალიზებელ ტექსტად წარმოდგენილია ერთი მთლიანი წინადადება. - მიუხედავად ამისა, სისტემამ სამ ერთმანეთისაგან საკმაოდ დაშორებულ ადგილას სამი ერთმანეთისგან განსხვავებული სახის სინტაქსური შეცდომა გამოიცნო.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ არცერთ ბიჭს, რომელსაც უყვარს ჭადრაკი, არ უყვარს ფეხბურთი. თუ ყველა მდინარე, რომელიც ტბაში, რომელიც ჩაედინება, ცივია, მაშინ ან არცერთი ცივი მდინარე ტბაში არ ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია. თუ პიანისტი შეასრულებს დიდ ფუგას, მაშინ ის არ შეასრულებს ზოგიერთ პატარა სონატას.

სურ. 34

სისტემამ შეამჩნია ზმნურ სახელურ ფრაზაში დაშვებული სინტაქსური შეცდომა.

ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ არცერთ ბიჭს, რომელსაც უყვარს ჭადრაკი, არ უყვარს ფეხბურთი. თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან არცერთი ცივი მდინარე ტბაში არ ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია. თუ პიანისტი შეასრულებს დიდ ფუგას, მაშინ ის არ შეასრულებს ზოგიერთ პატარა სონატას.

სურ. 35

სისტემამ გამოიცნო პირველ წინადადებაში დამწყები 'თუ' კავშირის ნაკლულობა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი მაშინ არცერთ ბიჭს, რომელსაც უყვარს ჭადრაკი, არ უყვარს ფეხბურთი. თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან არცერთი ცივი მდინარე ტბაში არ ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია. თუ პიანისტი შეასრულე დიდი ფუგა, მაშინ ის არ შეასრულებს ზოგიერთ პატარა სონატას.

სურ. 36

აქ სისტემამ პირველ წინადადებაში შეამჩნია მძიმის ნაკლულობა, მესამეში - სინტაქსური შეუთანხმებლობა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ არცერთ ბიჭს, რომელსაც უყვარს ჭადრაკი, არ უყვარს ფეხბურთი. თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, ან არცერთი ცივი მდინარე ტბაში არ ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია. თუ შენ შეასრულე დიდი ფუგა, მაშინ ის არ შეასრულებს ზოგიერთ პატარა სონატას.

სურ. 37

აქ სისტემამ მეორე წინადადებაში დაიჭირა 'თუ'-ს მეწყვილე 'მაშინ' სიტყვის ნაკლულობა.

თუ ყველა ბიჭს უყვარს ფეხბურთი მაშინ ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ არცერთ ბიჭს, რომელსაც უყვარს ჭადრაკი, არ უყვარს ფეხბურთი. თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან არცერთი ცივი მდინარე ტბაში არ ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია. შენ შეასრულე ფუგა, მაშინ ის არ შეასრულებს ზოგიერთ პატარა სონატას.

სურ. 38

სისტემამ პირველ და მესამე წინადადებებში დაიჭირა შეცდომები.

თუ ყველა ბიჭს უყვარს ფეხბურთი მაშინ ზოგიერთ ბიჭს უყვარს ჭადრაკი. არცერთ ბიჭს, რომელსაც უყვარს ჭადრაკი, არ უყვარს ფეხბურთი. ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია. ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივი. შენ შეასრულე ფუგა. ის შეასრულებს პატარა სონატას.

სურ. 39

აქ თავდაპირველი ტექსტი ექვს შედარებით მარტივ წინადადებად დავახურდავით. სისტემამ პირველ და მესამე წინადადებებში დაიჭირა შეცდომები.

ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს
ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს
ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება,
ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან
ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 40

სისტემამ ამ ერთ დიდ რთულ წინადადებაში დაადასტურა პირველი ქვეწინადადების
სინტაქსური გამართულობა და, ასევე, დაადასტურა მისი მეწყვილე ქვეწინადადების სინტაქსური
გაუმართაობა.

თუ ყველა ბიჭი უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს
ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს
ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება,
ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან
ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 41

სისტემამ დაიჭირა რთული წინადადების ქვეწინადადებაში სინტაქსური შეცდომა.

თუ ზოგიერთი ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს
ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს
ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება,
ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან
ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 42

სისტემამ დაიჭირა რთული წინადადების ქვეწინადადების სახელურ ფრაზაში სინტაქსური
შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 43

სისტემამ დაიჭირა რთული წინადადების 'თუ' ნაწილში სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკს, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 44

სისტემამ დაიჭირა რთული წინადადების ქვეწინადადებაში სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 45

სისტემამ დაიჭირა პირველი 'მაშინ' ოპერატორის წინა (მარცხენა) მძიმის ნაკლულობა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 46

სისტემამ შეამჩნია, რომ პირველ 'თუ'-ს მისი მეწყვილე 'მაშინ' აკლია.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ 47

სისტემამ დაიჭირა 'რომელსაც' ოპერატორის წინა (მარცხენა) მძიმის ნაკლულობა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭი, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ 48

სისტემამ დაიჭირა შეყვითლებულ არეში არსებული სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭი, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 49

სისტემამ დაიჭირა შეყვითლებულ არეში არსებული სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭი, რომელსაც გიყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 50

სისტემამ დაიჭირა შეყვითლებულ არეში არსებული სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭი, რომელიც გიყვარს შენ, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 51

სისტემამ დაიჭირა შეყვითლებულ არეში არსებული სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 52

სისტემამ დაიჭირა შეყვითლებულ არეში 'თუ'-ს არასინტაქსური განლაგება.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარემ, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 53

სისტემამ დაიჭირა შეყვითლებულ არეში არსებული სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბას ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 54

სისტემამ დაიჭირა შეყვითლებულ არეში არსებული სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 55

სისტემამ დაიჭირა შეყვითლებულ არეში 'მაშინ' სიტყვის წინა (მარცხენა) მძიმის ნაკლულობა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივმა მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 56

სისტემამ დაიჭირა შეყვითლებულ არეში არსებულ სახელურ ფრაზაში სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 57

სისტემამ დაიჭირა შეყვითლებულ არეში 'ან'-ის მეწყვილე 'ან'-ის ნაკლულობა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება და ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 58

სისტემამ დაიჭირა შეყვითლებულ არეში 'ან'-ის მეწყვილე 'ან'-ის ნაკლულობა, მიუხედავად იმისა, რომ იმ ადგილზე იჯდა კავშირი 'და'.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში არ ჩაედინება, ცივია.

სურ. 59

სისტემამ დაიჭირა შეყვითლებულ არეში არსებული სინტაქსური შეცდომა.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში, რომელსაც ცივია, არ ჩაედინება, ცივია.

სურ. 60

სისტემამ დაიჭირა ამ ერთ დიდ რთულ წინადადებაში არსებული სინტაქსური შეცდომა. - ამ წინადადების სინტაქსურად გამართული ფორმა შეგიძლიათ ნახოთ სურ. 61-ზე.

თუ ყველა ბიჭს უყვარს ფეხბურთი და ზოგიერთ ბიჭს უყვარს ჭადრაკი, მაშინ ყველა ბიჭს, რომელსაც უყვარს ჭადრაკი, უყვარს ფეხბურთი და თუ ყველა მდინარე, რომელიც ტბაში ჩაედინება, ცივია, მაშინ ან ყველა ცივი მდინარე ტბაში ჩაედინება ან ზოგიერთი მდინარე, რომელიც ტბაში, რომელიც ცივია, არ ჩაედინება, ცივია.

სურ. 61

სურ. 57-ზე მოცემული წინადადების სინტაქსურად გამართული ფორმა.

3.1.4. GeoNetSintaxAnalizer_2011 სისტემის შესაძლო გამოყენების არეებისა და განვითარების პერსპექტივების მიმოხილვა.

პროექტის ფარგლებში შესრულებული და ზემოთ მოკლედ აღწერილი GeoNetSintaxAnalizer_2011 სისტემა, რომელიც, ამავედროულად, არის ამ მიმართულებით წარმოებული მრავალწლიანი თეორიული და ექსპერიმენტული კვლევების შედეგი, ცალსახა დასტურია მიღწეული შედეგის შემდგომი გაღრმავებისა და გასრულების მიზნით კვლევების განვითარების აუცილებლობისა.²⁴

ამასთან, ამ უკვე არსებული GeoNetSintaxAnalizer_2011 სისტემის შემდგომი უახლოესი განვითარების პერსპექტივები და შესაძლო გამოყენების არეები შემდეგია:

1. უკვე არსებული GeoNetSintaxAnalizer_2011 სისტემა, მისი უბრალო ლექსიკური გასრულებითა და გამოყენებითი მხარეების მოხერხებულობაზე ორიენტირებული მარტივი პროგრამული ინსტრუმენტების დამატებით, დღესვე შეიძლება იქნეს გამოყენებული ქართული ენის სწავლისა და სწავლების მეტად მნიშვნელოვან ინსტრუმენტად. ამის გათვალისწინებით, უახლოეს მომავალში, ამ სისტემების მომხმარებლებად შეიძლება იქცნენ:

1. განათლებისა და მეცნიერების სამინისტრო, რათა ახალ საფეხურზე აიყვანოს და, შესაბამისად, ხარისხობრივად გააუმჯობესოს:

1. არაქართულენობრივ რეგიონებში ქართული ენის სწავლისა და სწავლების დღეს ქვეყანაში მიზნობრივად და ინტენსიურად მიმდინარე პროცესი;
2. დღეს უკვე თითქმის სრულად გათანამედროვეებულ - კომპიუტიზირებულ ქართულ სასკოლო-საგანმანათლებლო სფეროში თანამედროვე ტექნოლოგიების გამოყენებით ქართული ენის სწავლისა და სწავლების პროცესი;

²⁴ ეს ასეა, რადგან სრულიად ნათელია სრული (ორთოგრაფიულ-მორფოლოგიური, სინტაქსური, ლოგიკურ-სემანტიკური) ქართული მართლმწერი სისტემის რაც შეიძლება დროული შემუშავების აუცილებლობა!

3. მოზარდებისათვის ინტერნეტში სასწავლო-საგანმანათლებლო პროგრამების ინტენსიური და მიზნობრივი მიწოდების დღეს მიმდინარე პროცესი;
 2. **დიასპორის საკითხებში საქართველოს სახელმწიფო მინისტრის აპარატის ვებ-გვერდი „ქართული დიასპორა“.** - სისტემის ამ ვებ-გვერდზე განთავსება, ამუშავება და მისი თანდათანობითი განახლება-გაუმჯობესება, ცხადია, ხელს შეუწყობს ქართული დიასპორის ქართულ ენაში და, შესაბამისად, ქართველ ერთან უფრო მჭიდრო კულტურულ-ეკონომიკური ინტეგრირების დღეს მიზნობრივად მიმდინარე პროცესს;
 3. **ნებისმიერი კერძო პირი**, ვინც ფლობს და ამა თუ იმ დატვირთვით მოიხმარს ქართულ ენას ან სწავლობს მას იმ მიზნით, რომ შემდგომში ამა თუ იმ დატვირთვით მოიხმაროს იგი;
 4. **ნებისმიერი დაწესებულება**, რომელიც ამა თუ იმ დატვირთვით მოიხმარს ქართულ ენას.
2. **GeoNetSintaxAnalyzer_2011** სისტემა, რომელიც, როგორც უკვე აღინიშნა, GeoSintaxAnalyzer/Translator_2009 სისტემის GeoSintaxAnalyzer_2009 ქვესისტემის ASP.NET რეალიზაციისა და მისი შემდგომი გაფართოების შედეგია და რომელიც, ამავდროულად, მსგავსად ModusPonens_2005 სისტემისა, ქართულ ენობრივ ტექსტს აანალიზებს ამ ტექსტის მათემატიკური ენაზე წინასწარი დაყვანის ანუ გადათარგმნის საფუძველზე, უკვე გათვალისწინებულია მისი შემდგომი მნიშვნელოვანი განვითარების პერსპექტივები:²⁵
1. **GeoNetSintaxAnalyzer_2011** სისტემის განვითარება და გასრულება GeoSintaxAnalyzer/Translator_2009 სისტემის მიმართულებით, რაც ერთ მთლიან სისტემაში ინტეგრირებული სახით მოგვცემს ქართულ მრავალენოვან მთარგნელ სისტემას, რომელიც მათემატიკურ ენაში საშუაშაგლო თარგმანის მეშვეობით ქართულიდან/ქართულზე სხვადასხვა ენებში/ენებიდან გადათარგმნის/გადმოთარგმნის როგორც მარტივი, ისე რთული წინადადებებით აგებულ სხვადასხვა ლოგიკური სტრუქტურის მქონე ტექსტებს არა ალბათური, არამედ მკაცრი მათემატიკური სიზუსტით;
 2. **GeoNetSintaxAnalyzer_2011** სისტემის განვითარება და გასრულება ModusPonens_2005 სისტემის მიმართულებით, რაც ერთ მთლიან სისტემაში ინტეგრირებული სახით მოგვცემს ქართული ლოგიკური ამოცანების ამომხსნელ და შემმოწმებელ თითქმის სრულ ინტერაქტიურ ინტელექტუალურ სისტემას, რომელსაც გარდა ქართული ლოგიკური ამოცანების ამომხსნის და უკვე ამომხსნილი ამოცანების სისწორის შემოწმების უნარებისა, ექნება აგრეთვე უნარი მომხმარებელს დასაბუთებულად მიაწოდოს მოსაზრება ან იმის თაობაზე, თუ რატომ წარმართა ამოცანის ამომხსნა ისე, როგორც წარმართა, ანდა იმის თაობაზე თუ რატომ შეაფასა მომხმარებლის

²⁵ უნდა ხაზგასმით აღინიშნოს, რომ ამ აქ ხაზგასმული ზოგადი თვალსაზრისებით GeoNetSintaxAnalyzer_2011 სისტემაში უკვე რეალიზებული შესაძლებლობები ბევრად მეტია GeoSintaxAnalyzer/Translator_2009 და ModusPonens_2005 სისტემების ზოგად შესაძლებლობებზე.

მხრიდან შესასწორებლად შემოსული ამოცანის ამოხსნა ისე როგორც შეაფასა.

და ბოლოს, ცალკე უნდა გამოიყოს GeoNetSyntaxAnalyzer_2011 სისტემის გამოყენების კიდევ ერთი მხარე. ეს უკავშირდება უკვე შემუშავებულ Geotextreader_2011 და DemoGeotextreader_2011 მკითხველ სისტემებთან და შემუშავების პროცესში მყოფ ალოფონურ სინთეზატორთან მისი ინტეგრირებით განსაზღვრულ პერსპექტივებს.

კერძოდ, მოგახსენებთ, Geotextreader_2011 და DemoGeotextreader_2011 მკითხველი სისტემების მაგალითზე: როდესაც ჩვენ ვამბობთ, რომ Geotextreader_2011 და DemoGeotextreader_2011 მკითხველ სისტემებში რეალიზებულია მათი ამგები მეთოდის ჯერ მხოლოდ 5-10%, ჩვენ მხედველობაში გვაქვს ის, რომ ბევრად უფრო ადვილია სარეალიზაციო სისტემის, ვთქვათ ასე, 85%-იანი რეალიზაციის მიღწევა, ვიდრე ამ 85%-დან შედეგის 100%-იან რეალიზაციამდე აწევა. - ეს გასაგებია.

ახლა რა შუაშია აქ GeoNetSyntaxAnalyzer_2011 სისტემა. - საქმე ისაა, რომ ჩვენ დაგეგმილი გვაქვს გამოვიყენოთ ამ სისტემის უნარი გააკეთოს ტექსტის სრული სინტაქსური ანალიზი Geotextreader_2011 და DemoGeotextreader_2011 მკითხველ სისტემებში უკვე რეალიზებული არასემანტიკური ანუ გაუაზრებელი კითხვის სემანტიკურ ანუ გააზრებულ კითხვაზე გადასაცვანად,²⁶ რაც, როგორც ჩვენ ვთვლით (კ.ფხაკაძე, გ.ჩიჩუა, მ.ჩიქვინიძე), უკვე მიღწეულ კითხვის ხარისხის გარჩევადობისა და ბუნებრიობის თვალსაზრისით მაქსიმალურთან სულ ცოტა 40-60% პროცენტით მიაახლოებს.²⁷ - ამგვარად, GeoNetSyntaxAnalyzer_2011 სისტემის უნარი, ქართული ტექსტის გაანალიზებისას გამოარჩიოს და ცალკე გამოყოს როგორც მარტივი წინადადებების ამგები ზმნები და ასეთი წინადადებების შემადგენელი სხვადასხვა სახის სახელურ ფრაზები (იქნება ეს მარტივი სახელური ფრაზა, ზედსართაული სახელური ფრაზა, განზოგადებული სახელური ფრაზა, ზმნური სახელური ფრაზა), ასევე რთული წინადადებების ამგები ლოგიკური კავშირები და ასეთ წინადადებებში არსებული სასვენი ნიშნები მათი სხვადასხვა სახის ფუნქციონირების გათვალისწინებით, არის ის, რაც საშუალებას მოგვცემს Geotextreader_2011 და DemoGeotextreader_2011 მკითხველ სისტემებში უკვე რეალიზებული არასემანტიკური ანუ გაუაზრებელი კითხვა თითქმის ამომწურავი მოცულობით გადავიყვანოთ სემანტიკურ ანუ გააზრებულ კითხვაზე, რაც, ცხადია, საგრძნობლივ აწევს ამ სისტემებით უკვე მიღწეულ არადაბალ კითხვის ხარისხს.

4 ბოლოთქმის მაგიერ

პროექტით მიღწეული და ზემოთ საანგარიშოდ წარმოდგენილი შედეგები, ამავდროულად, არის შედეგი იმ მრავალწლიანი კვლევითი საქმიანობისა, რომელიც 2002 წელს სახელმწიფო-მიზნობრივ საფეხურზე გადავიდა თსუ სახელმწიფო-მიზნობრივი პროგრამის „კომპიუტერის სრული პროგრამულ-მომსახურებითი

²⁶ კითხვის ხარისხის აწევაში კითხვის ბუნებრივი გამოთქმის ანუ სემანტიკური კითხვის მნიშვნელობაზე იხილეთ [8].

²⁷ ანუ, ჩვენთან უკვე მიღწეულ ასე 85% -იან კითხვის ხარისხს სულ ცოტა აწევს 91-97%-იან კითხვის ხარისხამდე.

მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში” სახით²⁸ და რომელიც 4 წლიანი შეფერხების შემდეგ კვლავ ფეხზე დადგა საქართველოს ტექნიკური უნივერსიტეტის რექტორის - აკადემიკოს არჩილ ფრანგიშვილის ინიციატივის, აკადემიური საბჭოს 2010 წლის 24 დეკემბრის დადგენილების და წარმომადგენლობითი საბჭოს 29 დეკემბრის გადაწყვეტილების შესაბამისად „ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის“ ფორმირებით. - ის შედეგები, რომლებიც საქართველოს ტექნიკური უნივერსიტეტის მიერ დაფინანსებული პროექტით იქნა მიღებული, ცხადია არ იქნებოდა, რომ არა უნივერსიტეტის რექტორისა და მმართველი გუნდის ეს დროული მხარდაჭერა. - ამან მოგვცა ქართული ენის ტექნოლოგიების მიმართულებით კიდევ ერთი მნიშვნელოვანი ნაბიჯის გადადგმის საშუალება, რაც, ამავდროულად, არის ნაბიჯი მიზნით - დროულად შეიქმნას ქართული „კონკურენტუნარიანი საინფორმაციო გარემო“, რის გარეშეც, მომავალ თაობას, ცხადია, ძლიერ გაუჭირდება „დამოუკიდებლად გასცეს პასუხი დროის გამოწვევებს.“

„ამაშია საქართველოს გადარჩენა.“

მეცნიერებათა
ლოქტორი

/კ.ფხაკაძე /

სტუ ქართული ენის ტექნოლოგიების ცენტრის ხელმძღვანელი, თსუ ქართული ენის მათემატიკური დამუშავების მთავარი სპეციალისტი.

5. დამოწმებული ლიტერატურის ნუსხა:

1. Pkhakadze K., Abzianidze L., Maskharashvili A., Pkhakadze N., Chiqvinidze M., The Core Part of the Georgian Language As a Mathematical Theory, Book of Abstracts of II International Conference of Georgian Mathematical Union, pp.66, 2011;
2. კ.ფხაკაძე, ლ.აბზიანიძე, ა.მასხარაშვილი, ნ.ფხაკაძე, მ.ჩიქვინიძე, ქართული ენის ლოგიკური გრამატიკა და ქართულენოვანი კომპიუტერი, არნ.ჩიქობავას ენათმეცნიერების ინსტიტუტის II საერთაშორისო კონფერენციის „ქართული ენა და თანამედროვე ტექნოლოგიები“, გვ.132–133, 2011;
3. Abzianidze L., An HPSG-based Formal Grammar of a Core Fragment of Georgian Implemented in TRALE, Master Thesis, Charles University in Prague, pp.1-124, 2011;
4. მ.ჩიქვინიძე, თსუ ზუსტ და საბუნებისმეტყველო მეცნიერებათა ფაკულტეტი, საბაკალავრო თემა – „მათემატიკური მეთოდებით აგებული ქართულ-ინგლისურ-

²⁸ 2002 წელს, ამ მიზნობრივ პროგრამაზე, აკადემიკოსი ნოდარ ამაღლობელი წერდა:

„პროექტს დიდი სახელმწიფო მხარდაჭერა აქვს. მისი აქტუალობა დღეს არავითარ ეჭვს არ იწვევს, რადგან ეს არის ქართული ენის, როგორც კულტურული ფენომენის, შენარჩუნება-გადარჩენის, მისი განვითარებისა და შესწავლის გადრეკილვის საშუალება, ეროვნული ცნობიერების განმტკიცების ერთ-ერთი უმნიშვნელოვანესი გზა.

ამ ჭეშმარიტად ეროვნულ საქმეს საერთო მხარდაჭერა და თანადგომა სჭირდება“

- გერმანული ორმხრივ მთარგმნელი კომპიუტერული სისტემის 1 ვერსია“, თემას ხელმძღვანელობდა მენიერებათა დოქტორი კ.ფხაკაძე, გვ.1-50, 2010;
5. ა.მასხარაშვილი, თსუ ზუსტ და საბუნებისმეტყველო მეცნიერებათა ფაკულტეტი, სამაგისტრო თემა „ქართული მოქმედ პირიანი ზმნის მათემატიკური ანალიზი და ქართულ წერით ენაში ინტელექტუალური კომპიუტერული სისტემის პირველი ვერსია“, თემას ხელმძღვანელობდა მენიერებათა დოქტორი კ.ფხაკაძე, გვ.1-58, 2010;
 6. Pkhakadze K., Chichua G., Vashalomidze A., Gabunia K., Abzianidze L., Maskharashvili A., Pkhakadze N., Chiqvinidze M., The Grounding Questions of The Mathematical Theory of The Georgian Language and Thinking and Some Subsystems of The 1st Version of the Voice Managed Georgian Intellectual Computer System, Book of Abstracts of International Conference of Georgian Mathematical Union, pp.50-51, 2010;
 7. შ. ფხაკაძე, მათემატიკური ლოგიკა - საფუძვლები, III ნაწილი, „უნივერსალი“, გვ.1-192, 2009;
 8. Taylor P., Text-to-Speech Synthesis, Cambridge University Press, pp.627, 2009;
 9. კ.ფხაკაძე, გ.ჩიჩუა, ა.ვაშალომიძე, კ.გაბუნია, ლ.აბზიანიძე, ს.მასხარაშვილი, მ.ჩიქვინიძე, ქართული ენის და აზროვნების მათემატიკური თეორიის სასაფუძვლო საკითხები და ხმით მართვადი ქართული ინტელექტუალური კომპიუტერული სისტემის I ვერსიის ნაწილები, არნ.ჩიქობავას ენათმეცნიერების ინსტიტუტის პირველი საერთაშორისო კონფერენციის „ქართული ენა და თანამედროვე ტექნოლოგიები“ მასალები, გვ.60-66, 2009;
 10. კ.ფხაკაძე, ა.ვაშალომიძე, გ.ჩიჩუა, ლ.აბზიანიძე, ა.მასხარაშვილი, მ.ჩიქვინიძე, კ.გაბუნია, ქართული ენის და აზროვნების მათემატიკური თეორიის და ქართული ინტელექტუალური კომპიუტერული სისტემის შემუშავების მიზნები და ქართული ენის წინაშე მდგარი საფრთხეები, საქართველოს საპატრიარქოს წმინდა ანდრია პირველწოდებულის სახელობის ქართული უნივერსიტეტის გამოცემა, გვ.1-24, 2009;
 11. Pkhakadze K., Chichua G., Vashalomidze A., Abzianidze L., Maskharashvili A., Chiqvinidze M., The Georgian Text Reader System with User's Possibility to Build in an Own Synthetic Voice, REPORTS of Enlarged Sessions of The Seminar of I.Vekua Institute of Applied Mathematics, Vol. 23, pp.105-110, 2009;
 12. Pkhakadze K., Chichua G., Vashalomidze A., Abzianidze L., Maskharashvili A., The Georgian Reader-Listener System with User's Possibility to Build in an Own Synthetic Voice, REPORTS of Enlarged Sessions of The Seminar of I. Vekua Institute of Applied Mathematics, Volume 23, pp.100-104, 2009;
 13. გ.ჩიჩუა, სადოქტორო თემა „ქართული მეტყველების სინთეზი და გამოცნობა“, თემას ხელმძღვანელობს მეცნიერებათა დოქტორი კ.ფხაკაძე, საქართველოს საპატრიარქოს წმინდა ანდრია პირველწოდებულის სახელობის ქართული უნივერსიტეტი, კომპიუტერული ტექნოლოგიები და მათემატიკური მოდელირება, გვ.1-220, 2009;
 14. 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, Volume 23, pp.87-93, 2009;
 15. Meurer P., A Computational Grammar for Georgian, Logic, Language, and Computation, Subseries of Lecture Notes in Computer Science, TbiLLC'2007, Springer, pp.1-15, 2009;
 16. ა.მასხარაშვილი, თსუ ზუსტ და საბუნებისმეტყველო მეცნიერებათა ფაკულტეტი, საბაკალავრო თემა – „ზოგიერთი რთული წინადადების მათემატიკური ანალიზი“, თემას ხელმძღვანელობდა მენიერებათა დოქტორი კ.ფხაკაძე, გვ.1-45, 2008;
 17. ლ.აბზიანიძე, თსუ ზუსტ და საბუნებისმეტყველო მეცნიერებათა ფაკულტეტი, საბაკალავრო თემა – „მათემატიკური მეთოდებით აგებული მართლმწერი ქართული მარტივი წინადადებებისათვის“, თემას ხელმძღვანელობდა მენიერებათა დოქტორი კ.ფხაკაძე, გვ.1-37, 2008;
 18. ნ.ფხაკაძე, თსუ ზუსტ და საბუნებისმეტყველო მეცნიერებათა ფაკულტეტი, საბაკალავრო თემა – „შემამოკლებელი სიმბოლოები ქართულ ენაში“, თემას ხელმძღვანელობდა მენიერებათა დოქტორი კ.ფხაკაძე, გვ.1-49, 2008;

19. კ.ფხაკაძე, კ.გაბუნია, გ.ჩიჩუა, ა.მასხარაშვილი, ლ.აბზიანიძე, ნ.ვახანია, ნ.ფხაკაძე, ბ.ჩიქვინიძე, ლ.გურასაშვილი, ნ.ლაბაძე, ქართული ინტელექტუალური კომპიუტერული სისტემის კონსტრუირების მიზნები და ქართული ენის წინაშე მდგარი საფრთხეები, არნ.ჩიქობავას სახელობის ენათმეცნიერების ინსტიტუტის VI რესპუბლიკური კონფერენციის „ბუნებრივ ენათა დამუშავება“, მასალები, თბილისი, გვ.23-24/გვ.33-34, 2008;
20. Pkhakadze K., Abzianidze L., Maskharashvili A., Georgian Language's Theses, Seminar of I. Vekua Institute of Applied Mathematics REPORTS, Vol. 34, pp. 108-121, 2008;
21. Pkhakadze K., Chichua G., Abzianidze L., Maskharashvili A., About 1-Stage Voice Managed Georgian Intellectual Computer System, Seminar of I. Vekua Institute of Applied Mathematics REPORTS, Vol. 34, pp. 96-107, 2008;
22. კ.ფხაკაძე, ვ.ფხაკაძე, ლ.აბზიანიძე, ნ.ვახანია, ნ.ლაბაძე, ა.მასხარაშვილი, ნ.ფხაკაძე, ბ.ჩიქვინიძე ქართული ენის ბუნებრივი ლოგიკის სასაფუძვლო საკითხები, არნ.ჩიქობავას ენათმეცნიერების ინსტიტუტის V რესპუბლიკური კონფერენციის „ბუნებრივ ენათა დამუშავება“ მასალები, გვ.18-21/გვ.42-44, 2007;
23. თ.ჩიჩუა, გ.ჩიჩუა, კ.ფხაკაძე, ლ.ალთუნაშვილი, ზ.ცირამუა, რ.ბახტაძე, კომპიუტერი ქართული ტექსტების წაკითხვას შეძლებს, არნ.ჩიქობავას ენათმეცნიერების ინსტიტუტის VI რესპუბლიკური კონფერენციის „ბუნებრივ ენათა დამუშავება“ მასალები, გვ.25-27/გვ.36-37, 2007;
24. კ.ფხაკაძე, ვ.ლევიაშვილი, ვ.ფხაკაძე, მორფემების მათემატიკური ფუნქციებისა და წინადადებებში სიტყვათშორისი ცარიელი ადგილების პრობლემა ქართული ენის მიხედვით, არნ.ჩიქობავას ენათმეცნიერების ინსტიტუტის V რესპუბლიკური კონფერენციის „ბუნებრივ ენათა დამუშავება“ მასალები, გვ.16-18/გვ.40-42, 2006;
25. კ.ფხაკაძე, გ.ჭანკვეტაძე, ი.ბერიაშვილი, ქართულ-გერმანული მთარგმნელი და I საფეხურის პროგრამული სინთეზატორი და ანალიზატორი ქართულისათვის, სამეცნიერო-საგანმანათლებლო ჟურნალი „ქართული ენა და ლოგიკა“, (N2), „უნივერსალი“, გვ.88-91, 2005;
26. კ.ფხაკაძე, ვ.ფხაკაძე, ლ.ლევიაშვილი, ქართული ენის ზოგადი კვანტორულ-პრედიკატული ანალიზი და ლოგიკური ამოცანების ამომხსნელი პირველი ექსპერიმენტული პროგრამა ქართულისათვის, სამეცნიერო-საგანმანათლებლო ჟურნალი „ქართული ენა და ლოგიკა“, (N2), „უნივერსალი“, გვ.82-87, 2005;
27. კ.ფხაკაძე, გლობალიზაცია, ქართული ენა და სახელმწიფო მიზნობრივი პროგრამა „კომპიუტერის სრული პროგრამულ-მომსახურეობითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“, სამეცნიერო-საგანმანათლებლო ჟურნალი „ქართული ენა და ლოგიკა“, (N2), „უნივერსალი“, გვ.1-11, 2005,
28. კ.ფხაკაძე, რ.სხირტლაძე, ქართული ენის სინტაქსური საფეხურის კომპიუტერული მართლმწერის ერთი ექსპერიმენტის შესახებ, სამეცნიერო-საგანმანათლებლო ჟურნალი „ქართული ენა და ლოგიკა“, (N2), „უნივერსალი“, გვ.62-65, 2005;
29. კ.ფხაკაძე, ნ.ბუაძე, „ნატგრა“ ტიპის ზმნური სემანტიკები და პრინციპული განსხვავებები ქართულ და ინდოევროპულ უღლებებს შორის, სასწავლო სამეცნიერო ჟურნალი „ქართული ენა და ლოგიკა“, (N2), თბილისი, „უნივერსალი“, გვ.66-81, 2005;
30. კ.ფხაკაძე, ლოგიკური ბრუნებისა და ლინგვისტური მიმართებების საკითხისათვის ქართულში, საგანმანათლებლო-სამეცნიერო ჟურნალი „ქართული ენა და ლოგიკა“, იანვარი-ივნისი (N1), თბილისი, გამომცემლობა „უნივერსალი“, გვ.19-77, 2005;
31. ტ.ფუტყარაძე, ქართული ენის ბუნების ამსახველი ლოგიკური პარადიგმის პრობლემისათვის, საგანმანათლებლო-სამეცნიერო ჟურნალი „ქართული ენა და ლოგიკა“, ივლისი-დეკემბერი, (N2), თბილისი, გამომცემლობა „უნივერსალი“, გვ.11-15, 2005;
32. კ.ფხაკაძე, წინარე ზმნური სემანტიკური ერთეულის ცნება, პირის ნიშანთა საკითხი, ინტეგრალური და არაინტეგრალური გაგების ზმნური სიტყვა-შინაარსები და ქართული ზმნის არასრული ანუ პირველი სემანტიკური კლასიფიკაცია, ჟურნალის „ქართული ენა და

- ლოგიკა“ ყოველწლიური დამატება „ძირითადი და დამხმარე სახელმძღვანელოები თანამედროვე მათემატიკურ ენათმეცნიერებაში“ (N1), „უნივერსალი“, გვ.72-152, 2004;
33. კ.ფხაკაძე, გ.ჭანკვეტაძე, ლ.ტიბუა, მ.ივანიშვილი, ლ.ლევიაშვილი, ე.სოსელია, ქართული ენის ბირთვული ნაწილის კომპიუტერული მოდელი და მარტივი წინადადების სინტაქსური სინთეზატორი და ანალიზატორი, არნ. ჩიქობავას ენათმეცნიერების ინსტიტუტის II რესპუბლიკური კონფერენციის „ბუნებრივ ენათა დამუშავება“ მასალები, გვ.38-40/გვ.47-48, 2004;
 34. კ.ფხაკაძე, ბუნებრივი ქართული ენობრივი სისტემის ერთი ფრაგმენტი როგორც სიმრავლურ თეორიული სისტემა, ქართულ-ევროპული ინსტიტუტი, სამეცნიერო ჟურნალი „ივერია“, ტ.X-XI, პარიზი, 104-116, 2003-2004;
 35. K. Pkhakadze, G.Chankvetadze, L.Tibua, M.Ivanishvili, L.Lekiasvili, E.Soselia, 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, CLLS, pp.129-137, 2003;
 36. კ.ფხაკაძე, გ.ჭანკვეტაძე, ლ.ტიბუა, მ.ივანიშვილი, ლ.ლევიაშვილი, ე.სოსელია, ენობრივ-აზროვნებითი გაგებადობის ცნებისა და მარტივი და ინტეგრალური გაგების ზმნური სიტყვების შესახებ ქართულში, არნ.ჩიქობავას ენათმეცნიერების ინსტიტუტის I რესპუბლიკური კონფერენციის „ბუნებრივ ენათა დამუშავება“ მასალები, გვ.37-39, 2003;
 37. თ.უთურგაძე, გრამატიკული კატეგორიებისა და მათი ურთიერთმიმართებისათვის ქართულში, „ქართული ენა“, თბილისი, გვ.1-244, 2002;
 38. კ.ფხაკაძე, სიტყვათა შორის სინტაქსური კავშირების შესახებ, ქართულ-ევროპული ინსტიტუტი, სამეცნიერო ჟურნალი „ივერია“, ტ. VII-VIII, პარიზი 2001 – 2002, 117-126;
 39. დ.მელიქიშვილი, ქართული ზმნის უღლების სისტემა, „ლოგოს პრესი“, თბილისი, გვ.362, 2001;
 40. Jurafsky D., Martin J.H., Speech and Language Processing, Prentice Hall, pp.934, 2000;
 41. შ. ფხაკაძე, მათემატიკური ლოგიკა - საფუძვლები, II ნაწილი, თსუ, გვ.1-267, 1999;
 42. ტ.ფუტკარაძე, ზმნის ფორმაწარმოების მოდელირებისათვის (კომპიუტერული მოდელები) - ენის კონცეპტუალური და კომპიუტერული მოდელები, III, "ქართული ენა", თბილისი, გვ.??-??, 1998;
 43. ლ.კვაჭაძე, თანამედროვე ქართული ენის სინტაქსი, თბილისი, გვ.1-554, 1996;
 44. შ. ფხაკაძე, მათემატიკური ლოგიკა - საფუძვლები, I ნაწილი, თსუ, გვ.1-280, 1996;
 45. Partee B.H., Meulen A., Wall E.R., Mathematical Methods in Linguistics, Kluwer Publisher, pp.663, 1990;
 46. Harris A.C., Georgian Syntax - A Study in Relational Grammar, Cambridge University Press, pp.327, 1981;
 47. ა.შანიძე, ქართული ენის გრამატიკის საფუძვლები, თსუ, 1980, გვ. 1-671, თბილისი;
 48. თ. გამყრელიძე, ზმნის „პირიანობა“ და „ვალენტობა“, გ.მაჭავარიანის ხსოვნასდმი მიძღვნილი საენათმეცნიერო კრებული „მეცნიერება“, თბილისი, გვ.33-50, 1979;
 49. Пхакадзе Ш.С., Некоторые Вопросы Теории Обозначений, ТГУ, с.1-195, 1977;
 50. Monague R., Formal Philosophy – Selected Papers of Richard Monague Edited and with an Introduction by Richard H.Thomason, Yale University, pp.1-369, 1974;
 51. არნ.ჩიქობავა, მარტივი წინადადების პრობლემა ქართულში, სსსრ მეცნიერებათა აკადემია – ენათმეცნიერების ინსტიტუტი, გვ.1-295, 1961;
 52. Успенский В.А., К определению падежа по А.Н.Колмогорову, Бюлетен Объединения по проблемам машинного перевода, 5, с.11-186 1957.

დანართი II

ალექსანდრე ვაშალომიძე

ქართული ტექსტის გახმოვანება

შინაარსი

1. შესავალი
2. სინთეზის მოდელი /II კვარტალი/
3. ტრიფონები. მსგავსი ტრიფონები. აკუსტიკური ბაზა /II კვარტალი/
4. კლასებად დაყოფის წესები /II კვარტალი/
5. ქართული ტექსტების სტატისტიკური ანალიზი.
წარმომადგენელთა ბაზა /III კვარტალი/
6. აკუსტიკური ბაზის ჩაწერა /III კვარტალი/
7. სიტყვების სეგმენტაცია აკუსტიკურ ერთეულებად /IV კვარტალი/
8. სინთეზი და ტესტირება /IV კვარტალი/
9. დასკვნები
10. ლიტერატურა

ქართული მეტყველების სინთეზატორების შექმნაზე მუშაობა საქართველოში დიდი ხანია მიმდინარეობს. შედეგად გვაქვს სხვადასხვა ავტორებისა და ჯგუფების მიერ შექმნილი სინთეზატორები. პროექტის პირველი პუნქტი სახელით "ქართული ტექსტების გახმოვანება" გულისხმობს ქართული მეტყველების სინთეზატორის შექმნას მეტყველების ბუნებრიობის შედარებით მაღალი ხარისხით.

ამ მიზნით პროექტის ფარგლებში განვახორციელეთ სხვადასხვა ხასიათის კვლევები და ჩავატარეთ მიღებული შედეგების ტესტირება. ქვემოთ მოკლედ წარმოვადგენთ ანგარიშს კვლევებისა და ტესტირების შედეგების შესახებ.

2. სინთეზის მოდელი

სინთეზის მოდელად არჩეული იყო სინთეზი ბუნებრივი ენის წინასწარ ჩაწერილი მონაკვეთებით, მინიმალურ მონაკვეთებად კი - ალოფონები. ასეთ სინთეზს ეწოდება ალოფონური. ის წარმოადგენს კონკატენაციური სინთეზის ეგრეთწოდებულ მეორე გენერაციას - სინთეზი ერთეულთა შერჩევის (**Unit Selection Synthesis**) გზით [1, 2, 3]. ასეთივე წესით აირჩევა ბაზა ეგრეთწოდებული კონტექსტ-დამოკიდებული (**Context-sensitive modelling**) მოდელის გამოყენების შემთხვევაშიც [1,2].

არჩევანის მიზეზი არის ის, რომ ალოფონური ბაზა არის უმცირესი ვარიანტი ყველა შესაძლო ვარიანტებს შორის (დიფონები, დიალოფონები, მარცვლები, ალომარცვლები, ნახევარ მარცვლები, ნახევარსიტყვები, სიტყვები, და ა. შ.). დიფონური ბაზის მოცულობაც არ არის დიდი, მაგრამ მისი გამოყენების დროს დგება დიფონთა შეერთების პრობლემა.

3. ტრიფონები. მსგავსი ტრიფონები. აკუსტიკური ბაზა.

მეტყველების პროცესში ალოფონის კონკრეტულ ჟღერადობაზე გავლენას ახდენს მრავალი ფაქტორი, შესაბამისად, რეალურ ალოფონთა რაოდენობა ძალიან დიდია. სინთეზატორში ალოფონთა ბაზის შემცირების მიზნით ჩვენც შემოვიფარგლეთ მხოლოდ იმ გავლენებით, რომლებსაც ალოფონზე ახდენენ მისი უშუალო მარცხენა და მარჯვენა მეზობლები; ასევე, გავითვალისწინეთ ალოფონის პოზიცია სიტყვის თავკიდურ, შუა და ბოლოკლიდურ პოზიციებში. ცხადია, ალოფონთა შერჩეული ვარიანტები ვერ ასახავს ალოფონთა მთლიან სპექტრს, მაგრამ, ვფიქრობთ, საკმარისად წარმოჩენს მოცემულ ფონემას.

სინთეზატორების შინაგანი სტრუქტურა შედგება ორი ნაწილისაგან. ერთეულთაგან შემდგომი „წინა“ და „უკანა“ ნაწილებისაგან (front-end, back-end). სხვანაირად მათ მოიხსენიებენ როგორც ტექსტური პროცესორი და მეტყველების პროცესორი.

ტექსტური პროცესორის დანიშნულებაა მიიღოს ტექსტის ნაკადი, გადაიყვანოს არაკითხვადი (სპეციალური) სიმბოლოები (ციფრები, თარიღები, პროცენტი, ნომერი, პარაგრაფი, ფულადი ერთეული და სხვა) ლექსიკურ წარმოდგენაში. ამ პროცესს ეწოდება ტექსტის ნორმალიზაცია (ტოკენიზაცია, პრე-პროცესორი). ნორმალიზებული ტექსტი იშლება წინადადებებად და წინადადებები - სიტყვებად. შემდეგ ყოველი

სიტყვა უნდა წარმოადგეს ფონეტიკური ალფაბეტის მიმდევრობის სახით (ფონეტიკური ტრანსკრიფცია, ტექსტი-ფონემაში, გრაფემა-ფონემაში). ეს იფორმაცია, პროსოდიკულ მახასიათებლებთან ერთად და გადაეცემა მეორე ნაწილს - მეტყველების პროცესორს, რომელიც მიღებულ ინფორმაციას ამეტყველებს. მეტყველების პროცესების აღწერისათვის გამოიყენება ოთხი ტერმინი: ფონემა, ფონი, ალოფონი და სეგმენტი [1].

გამარტივებული სქემა ასე გამოიყურება: **ტექსტის ნაკადი → ნორმალიზაცია → წინადადებები → სიტყვები → გრაფემა → ფონეტიკური ტრანსკრიფცია → ფონემა → ალოფონი → სეგმენტი → მეტყველება.**

ტექსტური პროცესორის ამოცანები, გამომდინარე ქართული ანბანის თავისებურებიდან გამარტივებულია. ქართულ ანბანში გვაქვს 33 სიმბოლო, რომელსაც შეესაბამება 33 სამეტყველო ბგერა /ფონემა/. მათ რიცხვს ემატება სიტყვათა შორის ინტერვალი - ' _ ', რომელიც სამეტყველო ნაკადში ისეთივე არსებით როლს ასრულებს, როგორც ჩვეულებრივი ფონემა; მაშასადამე, სულ განვიხილავთ 34 სიმბოლოს და მათ შესაბამის 34 სამეტყველო ბგერას /ფონემას/.

ტექსტის ნორმალიზაციის შემდეგ, ჩვენ ვრეზულობთ სიმბოლოთა მიმდევრობას, რომლებიც წარმოადგენენ ფონემათა სიმბოლურ აღნიშვნებს. (განსხვავებით სხვა ენებისაგან). ამიტომ ჩვენ შეგვიძლია გამოვტოვოთ ტექსტის (სიტყვების) ფონეტიკური ტრანსკრიფციით წარმოადგენის ეტაპი. ნორმალიზებული ტექსტის ნაკადის მიმდინარე სიმბოლოს სინთეზის პროცესი შეიძლება წარმოვადგინოთ ასე:

მიმდინარე სიმბოლო (გრაფემა) → ფონემა → მარცხენა და მარჯვენა მეზობლები → ალოფონი → სეგმენტი → მეტყველება.

ანუ ტექსტის მიმდინარე სიმბოლო განსაზღვრავს ფონემას, ხოლო მარცხენა და მარჯვენა მეზობლები - ალოფონს (ჩვენს მიერ მიღებული შეთანხმებით). სინთეზის პროცესის აღწერის გამარტივების მიზნით შემოვიღოთ ფორმალური აღნიშვნები.

S = " _ ა ბ გ დ ე ვ ზ თ ი კ ლ მ ნ ო პ ჟ რ ს ტ უ ფ ქ ჯ ყ მ რ ც ძ წ ჭ ხ ჯ ჳ " - ტექსტის მიმდინარე სიმბოლოთა სიმრავლე (ბაზაში გვაქვს მათი ალოფონები);

GL = " _ ა ბ გ დ ე ვ ზ თ ი კ ლ მ ნ ო პ ჟ რ ს ტ უ ფ ქ ჯ ყ მ რ ც ძ წ ჭ ხ ჯ ჳ " - ტექსტის მიმდინარე სიმბოლოს მარცხენა მეზობელთა სიმრავლე (მარცხენა კონტექსტები);

GR = " _ ა ბ გ დ ე ვ ზ თ ი კ ლ მ ნ ო პ ჟ რ ს ტ უ ფ ქ ჯ ყ მ რ ც ძ წ ჭ ხ ჯ ჳ . ! ? " - ტექსტის მიმდინარე სიმბოლოს მარჯვენა მეზობელთა სიმრავლე (მარჯვენა კონტექსტები);

ამ ეტაპზე ჩვენ განვიხილავთ **GR**-ის შემცირებულ ვარიანტს, რომელიც არ შეიცავს პუნქტუაციის სიმბოლოებს /მათი განხილვა განზრახული გვაქვს სამომავლოდ/.

ანბანის სიმბოლოთა დალაგებულ სამეულებს ეწოდებათ **ტრიფონები** [1,2,4] (ზოგჯერ ტრიფონები განისაზღვრებიან როგორც ფონემათა დალაგებული სამეულები. ეს გასაგების ენათა უმრავლესობისათვის, მაგრამ ვინაიდან ქართულში სიმბოლოთა მიმდევრობა და ფონემათა მიმდევრობა ურთიერთცალსახა თანადობაშია, ჩვენთვის მოსახერხებელია სიმბოლური ნოტაცია). ტრიფონთა სიმრავლე წარმოადგენს **GL, S, GR** სიმრავლეთა პირდაპირ ნამრავლს:

$$\text{TriPhone} = \text{GL} \times \text{S} \times \text{GR}.$$

ტრიფონები, მათი განმარტებიდან გამომდინარე, წარმოადგენენ ალოფონების კომბინატორულ და პოზიციურ (თავკიდურ, შუა და ბოლოკიდურ) ვარიანტებს. ქვემოთ ტექსტში ალოფონის ნაცვლად ზოგჯერ გამოყენებული იქნება "ტრიფონი" (ამ ტერმინით, როცა ეს არ იწვევს გაუგებრობას, აღნიშნულია როგორც სიმბოლოთა მომდევრობა, ასევე მისი შესაბამისი ალოფონიც).

მაგალითისათვის, სიტყვა "გამარჯობა" ტრიფონებით წარმოდგინება ასე:

"გამარჯობა" = "_გა" + "გამ" + "ამა" + "მარ" + "არჯ" + "რჯო" + "ჯობ" + "ობა" + "ბა_".

ყოველი ტრიფონი განსაზღვრავს ალოფონს (მის შეზღუდულ ვარიანტს). ასეთი წესით შესაძლებელია განვასხვაოთ ალოფონის როგორც კომბინატორული, ასევე პოზიციური ვარიანტები. ტრიფონთა ვარიანტული რაოდენობა ტოლია

$$|TriPhone| = |GL| \times |S| \times |GR| = 34 \times 34 \times 34 = 39304. \text{ (GR შემცირებულია)}$$

მათ შორის არ რეალიზდება ვარიანტები:

"_ _ _", "_ _ GR{ }", "GL{ } _ _". მათი რაოდენობაა $1 + 33 + 33 = 67$. საბოლოოდ თეორიულად საინტერესო ტრიფონთა რაოდენობაა **39237**.

ამ რაოდენობის შემდგომი შემცირების მიზნით შემოვიტანოთ სიმბოლოთა (ფონემათა) და ტრიფონთა (ალოფონთა) მსგავსების ცნებები.

რაიმე **s** სიმბოლოს შესაბამისი ტრიფონი აღვნიშნოთ - **xsy**, სადაც **x** არის მარცხენა, ხოლო **y** - მარჯვენა მეზობელი (აქცენტს ვაკეთებთ ფონემისა და ალოფონის სიმბოლოურ აღნიშვნაზე, ვინაიდან სინთეზის დროს საჭირო ფონემისა და ალოფონის იდენტიფიკაცია ხდება სწორედ ტექსტის სამელემენტიანი მონაკვეთებით, ანუ ტრიფონებით). **S** სიმრავლის ყოველი სიმბოლოსათვის /გარდა'_'_ სიმბოლოსი/, გვაქვს $34 \times 34 = 1156$ ტრიფონი. '_'_ სიმბოლოსათვის - $33 \times 33 = 1089$ ტრიფონი.

მარცხენა და მარჯვენა მეზობელთა სიმრავლეები **GL** და **GR** (შემცირებული), რაიმე წესით დავყოთ თანაუკვეთ კლასებად: მარცხენა მეზობელთა კლასები და მარჯვენა მეზობელთა კლასები. აღვნიშნოთ ისინი შესაბამისად **LC1, LC2, ... , LCm** და **RC1, RC2, ... RCn**.

საზოგადოდ, მარცხენა კლასების სიმრავლე განსხვავდება მარჯვენა კლასების სიმრავლისაგან.

ორ სიმბოლოს (შესაბამის ორ ფონემას) მარცხენა ან მარჯვენა მეზობელთა სიმრავლიდან ვუწოდოთ მსგავსი, თუ ისინი ეკუთვნიან ერთ მარცხენა ან მარჯვენა კლასს.

თუ ისინი ეკუთვნიან მარცხენა კლასს, ვუწოდოთ მარცხენა მსგავსი, თუ მარჯვენა კლასს - მარჯვენა მსგავსი.

სიმბოლოთა (ფონემათა) მსგავსება უნდა გავიგოთ ასე: ისინი ახდენენ იდენტურ გავლენას მისი მარჯვენა (მარცხენა) მეზობლის აკუსტიკურ ჟღერადობაზე.

ორ ტრიფონს (შესაბამის ორ ალოფონს) **x1s1y1** და **x2s2y2** ვუწოდოთ ტოლი, თუ **x1 = x2**, **s1 = s2** და **y1 = y2**.

ორ ტრიფონს (შესაბამის ორ ალოფონს), $x_1s_1y_1$ და $x_2s_2y_2$ ვუწოდოთ მსგავსი, თუ $s_1=s_2$ და სრულდება პირობები: x_1 და x_2 არიან მარცხენა, ხოლო y_1 და y_2 მარჯვენა მსგავსი სიმბოლოები.

სხვა სიტყვებით რომ ვთქვათ, მარცხენა კლასებში - **CL** გაერთიანებულნი არიან ის სიმბოლოები /ფონემები/, რომლებიც მის მარჯვნივ მდგომი ფონემის აკუსტიკურ ჟღერადობაზე ახდენენ იდენტურ გავლენას. ასევე მარჯვენა კლასებში - **CR** გაერთიანებულნი არიან ის სიმბოლოები, რომლებიც მის მარცხნივ მდგომ ფონემის აკუსტიკურ ჟღერადობაზე ახდენენ იდენტურ გავლენას.

აქედან გამომდინარე, ორი მსგავსი ტრიფონი (მათი შესაბამისი ორი ალოფონი), მსგავსების განმარტებიდან გამომდინარე, **მსგავსია აკუსტიკურად**. ანუ ისინი სმენის ორგანოთა მიერ აღიქმებიან, როგორც პერცეფციულად ერთმანეთთან ახლოს მდგომნი.

ტრიფონთა (ალოფონთა) მსგავსების მიმართება არის ექვივალენტობის მიმართება. ამიტომ **TriPhone** სიმრავლე დაიყოფა მსგავსთა თანაუკვეთ კლასებად. მათი რაოდენობა იქნება $m \times |S| \times n$ - მარცხენა კლასების რაოდენობა გამრავლებული ანბანის სიმბოლოთა რაოდენობაზე და გამრავლებული მარჯვენა კლასების რაოდენობაზე.

მიღებული ფაქტორ-სიმრავლე აღვნიშნოთ - **LSR**-ით (მარცხენა კლასი, სიმბოლო, მარჯვენა კლასი). მიღებული სიმრავლის სიმძლავრეა $|LSR| = m \times |S| \times n$.

რაიმე წესით ავირჩიოთ თითოეული კლასიდან თითო წარმომადგენელი: lc_1, lc_2, \dots , rc_1, rc_2, \dots , rc_n . წარმომადგენელთა სიმრავლე აღვნიშნოთ **RepCL**. მისი სიმძლავრეც ტოლია $m \times |S| \times n$.

წარმომადგენელთა სიმრავლე RepCL არის სწორედ აკუსტიკურ ერთეულთა (ალოფონთა) საძიებელი მინიმალური ბაზა.

სიმრავლეს **TriPhone** ვუწოდოთ ზუსტ ალოფონთა ბაზა (ტრიფონის სიზუსტით); სიმრავლეს **RepCL** ვუწოდოთ წარმომადგენელთა ბაზა.

4. კლასებად დაყოფის წესები.

ზემოთ მოყვანილ აღწერებში, მარცხენა და მარჯვენა მეზობელთა სიმრავლეები დაყოფილია თანაუკვეთ ქვესიმრავლეებად "რაიმე წესით". ასევე "რაიმე წესით" არიან არჩეული წარმომადგენლები მსგავსების კლასებიდან. კონკრეტული ბაზის მისაღებად საჭიროა დავაკონკრეტოთ ეს ორი "რაიმე წესი" /ამ "წესების" სხვადასხვა გზით განსაზღვრა გვაძლევს მოცულობით და შემადგენლობით განსხვავებულ ბაზებს/.

მაგალითად, თუ მარცხენა და მარჯვენა კლასებს დავაჯგუფებთ ასე: **CL = CR = {'_'; მყდერები; ყრუები}**, მიღებული კლასების სიმძლავრე იქნება $3 \times 34 \times 3 - 67 = 239$. ასეთი დანაწილება მოყვანილია მაგალითისათვის.

ალოფონთა ბაზების განსაზღვრის ზემოთ მოყვანილი წესების გარდა შესაძლებელია (და რეალიზებულები სხვადასხვა სინთეზატორში) სხვა წესებიც, მაგალითად ხმოვნებისათვის და თანხმოვნებისათვის გამსხვავებული კლასები. ან რაიმე სხვა. მათ ნაწილობრივ განვიხილავთ ქვემოთ.

კლასებად დაყოფისთვის შესაძლებელია გამოვიყენოთ სამი მიდგომა:

- ა) არტიკულაციური;
- ბ) აკუსტიკური;
- გ) ნეირონული ქსელები.

არტიკულაციური მიდგომა ითვალისწინებს ფონოტაქტიკისა და კოარტიკულაციის წესებს, ბგერათა დაჯგუფებას წარმოშობის ადგილისა და რაგვარობის მიხედვით და სხვა.

აკუსტიკური მიდგომა ეყრდნობა ბგერათა აკუსტიკურ მახასიათებლებს.

ნეირონული ქსელების გამოყენებისას ხდება სიგნალთა (წინასწარი სათანადო გარდაქმნების შემდეგ) კლასიფიკაცია და კლასტერიზაცია ნეირონული ქსელების საშუალებით. ამისათვის არსებობს სტანდარტული საშუალებები(LVQ, კოხენერის და სხვა სახის ქსელები[5,6]).

ერთ კლასტერში დაჯგუფებული ვექტორები შეესაბამებიან "მსგავს" სიგნალებს, ხოლო კლასტერთა რაოდენობა - კლასების რაოდენობას. ხელოვნური ინტელექტის გამოყენება აუცილებელია, რადგან ალოფონთა რაოდენობა ძალიან დიდია და მათი შედარება "სიახლოვის" თვალსაზრისით მარტივი გზებით შეუძლებელია. ამავე დროს მიღებული კლასტერები შესაძლებელია გამოვიყენოთ როგორც მეტყველების სინთეზის, ასევე ამოცნობის ამოცანებისათვის.

საინტერესოა სამივე გზით მიღებული შედეგების შედარება და შეჯერება /რაც მომავალი კვლევის საგანია/.

წარმოდგენილ პროექტში ალოფონები დავაჯგუფეთ არტიკულაციური მიდგომით: ცალკე ჯგუფებად წარმოდგენილი ხმოვნებისა და სონორების გარდა 6 ჯგუფად დაიყო ჩქამიერი ფონემები; კერძოდ, ჰომოგენური ფონემები დაჯგუფდა ლოკალური რიგების მიხედვით, მაგ., ბაგისმიერი და უკანაენისმიერი /ე.წ. ლაბიოველარული ბგერები/ მჟღერები ერთ ჯგუფში გაერთიანდნენ, წინაენისმიერები - მეორეში; შესაბამისად, ორ-ორ ჯგუფებად დავყავით ყრუ-ფშვინვიერი და ყრუ-მკვეთრი ფონემებიც. სამეტყველო ბგერათა დაჯგუფების ამგვარი სქემა შემოთავაზებულია პროექტის მენეჯერის, ტარიელ ფუტკარაძის მიერ. მის მიერ შემოთავაზებულ ვარიანტში მარცხენა კლასების სიმრავლე განსაზღვრულია შემდეგი სახით:

1. $LC = \{ _ , 'ა', 'ე', 'ი', 'ო', 'უ', 'მწლ', 'დძჯზჟ', 'ვზგლ', 'სცმჩთ', 'პყკ', 'ფქჰჳ', 'ტწჭ' \}$.

ხოლო მარჯვენა კლასების სიმრავლე ასე:

2. $RC = \{ _ , 'ა', 'ე', 'ი', 'ო', 'უ', 'ვმზგლ', 'ნლრზყდძჯ', 'ფქჰჳთცჩსშ', 'პყკტწჭ' \}$.

გვაქვს $m = 13$ მარცხენა და $n = 10$ მარჯვენა კლასი. ტრიფონთა სიმრავლის სიმძლავრე ტოლია

$$|LSR| = m * |S| * n = 13 * 34 * 10 = 4420.$$

თეორიულად არასაინტერესოა ვარიანტები, სადაც ერთმანეთის მიყოლებით გვაქვს ორი ან მეტი " _ " სიმბოლო. მათი რაოდენობაა $1 + 12 + 9 = 22$. საჭირო კლასების რაოდენობაა $4420 - 22 = 4398$.

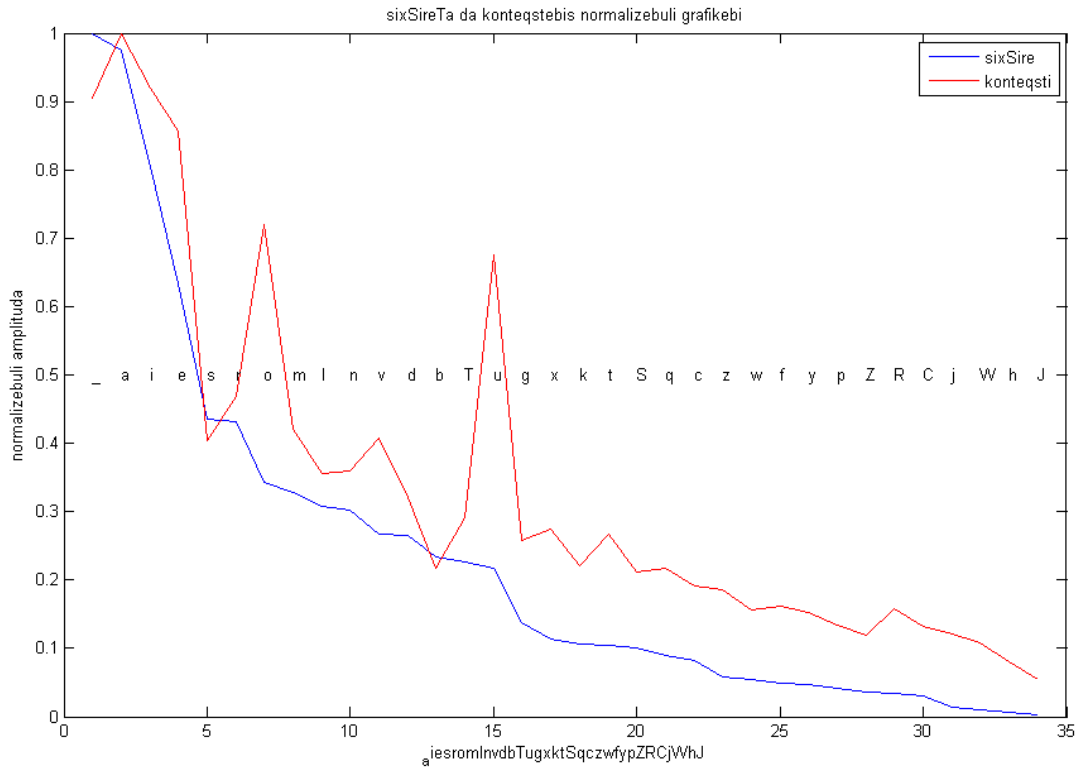
მორიგი ამოცანაა **4398** კლასიდან ავირჩიოთ თითო წარმომადგენელი. ამისათვის საჭიროა ქართული ტექსტების სტატისტიკური ანალიზი.

რაოდენობა, მესამეში პირველი ცხრილიდან გადმოტანილია შესაბამისი სიმბოლოს სიხშირე.

ცხრილი 2.

| სიმბოლო | რეალიზებული კონტექსტი (%) | სიხშირე |
|---------|---------------------------|---------|
| – | 69.5502 | 161185 |
| ა | 76.9031 | 157051 |
| ე | 65.9170 | 102 141 |
| ო | 70.7612 | 129 763 |
| ჟ | 55.3633 | 55 346 |
| ს | 51.9896 | 35 100 |
| ზ | 16.6955 | 37 647 |
| ბ | 19.8097 | 22 262 |
| ნ | 24.6540 | 42 866 |
| კ | 31.3149 | 42 949 |
| ღ | 14.2734 | 9 325 |
| თ | 22.4048 | 36 394 |
| კ | 16.9550 | 17 270 |
| ლ | 27.3356 | 49484 |
| მ | 32.3529 | 52 938 |
| ნ | 27.5952 | 48 699 |
| პ | 10.2941 | 66 29 |
| ჟ | 4.2388 | 455 |
| რ | 35.9862 | 69 639 |
| ს | 31.0554 | 70 107 |
| ტ | 20.5017 | 16 801 |
| ფ | 12.4567 | 8 022 |
| ქ | 16.6955 | 14 559 |
| ც | 12.1972 | 5 411 |
| ც | 11.7647 | 7 658 |
| შ | 16.2630 | 16 092 |
| ჩ | 10.2076 | 4 779 |
| ც | 14.7059 | 13 178 |
| ძ | 9.1696 | 5 887 |
| წ | 12.0242 | 8 885 |
| ჭ | 8.3045 | 1 656 |
| ხ | 21.1073 | 18 279 |
| ჯ | 9.3426 | 2 175 |
| ჰ | 6.1419 | 1 146 |

მეტი თვალსაჩინოებისათვის მეორე სურათზე მოცემულია სიხშირეებისა და რეალიზებული კონტექსტების ნორმალიზებული გრაფიკები.



სურ. 2. სიხშირეებისა და რეალიზებულ კონტექსტების ნორმალიზებული გრაფიკები.

სურათიდან ჩანს, რომ რეალიზებულ კონტექსტთა რაოდენობა საზოგადოდ კლებულობს სიმბოლოს სიხშირის კლებასთან ერთად. გვაქვს გამონაკლისებიც, ძირითადად ხმოვნების სახით. მაგალითად სიმბოლოები /ს/, /რ/, /მ/, /ლ/, /ნ/, /ვ/, /დ/, /ზ/, /თ/ უფრო ხშირია ტექსტში ვიდრე ხმოვანი /უ/ (ლურჯი გრაფიკი. ტექსტის ნაბეჭდ ვარიანტში გამოიყენეთ სურათის მარჯვენა ზედა კუთხეში მოთავსებული ლეგენდა), მაგრამ უკანასკნელის მიერ რეალიზებული კონტექსტების პროცენტული რაოდენობა მეტია მათ შესაბამის მახასიათებელზე (წითელი გრაფიკი).

ხმოვნების მიერ რეალიზებულ კონტექსტთა საშუალო მაჩვენებელია 64.1868%, ხოლო თანხმოვნების - 20.1928 %. ეს ნიშნავს, რომ ხმოვნები ყველა შესაძლო 1156 კონტექსტიდან საშუალოდ ახდენენ 64 %-ის რეალიზაციას, ხოლო თანხმოვნები - 20%-ს. მომავალში ეს შედეგი უნდა გავითვალისწინოთ და ხმოვნებისათვის და თანხმოვნებისათვის უნდა განვსაზღვროთ მსგავსების განსხვავებული კლასები. ასეთი პრაქტიკა გამოყენებულია მაგალითად ინგლისური და რუსულენოვანი სინთეზატორებისათვის ალოფონური ბაზების განსაზღვრის დროს [9,10].

კონკრეტული სიმბოლოს მიერ რეალიზებული ალოფონების ზუსტი სურათი მოცემულია ზემოთ აღწერილ კონტექტური მატრიცათა ვექტორის სახით. ის ჩაწერილია როგორც ექსელის 34 გვერდიანი ფაილი. ყოველ გვერდზე განთავსებულია შესაბამისი მატრიცა სახელით MATRIX_s, სადაც s აღნიშნავს შესაბამის სიმბოლოს. ფაილის სახელია CONTEXT_MATRIX.xlsx. სიდიდის გამო მისი აქ სრულად მოყვანა არ ხერხდება.

სურ. 3. CONTEXT_MATRIX.xlsx ფაილის MATRIX_a გვერდი.

დავუბრუნდეთ ჩვენს ძირითად მიზანს - მსგავსების კლასებიდან თითო წარმომადგენლის არჩევის საკითხს. ახლა ჩვენ ხელთ გვაქვს შესაბამისი მონაცემები. ისინი მოყვანილია CONTEXT_MATRIX.xlsx ცხრილებში. მაგალითად გვინდა ავირჩიოთ /ა/ სიმბოლოსათვის მსგავსების კლასების წარმომადგენლები. ამისათვის ფაილში CONTEXT_MATRIX.xlsx გავხსნათ გვერდი სახელით MATRIX_a (სურ. 3). ცხრილი დაყოფილია ჩარჩოებით მსგავსების კლასებად. ყოველი მსგავსების კლასი წარმოადგენს მინორს, რომლის სრტიქონისა და სვეტის თანაკვეთაზე დგას შესაბამისი კონტექსტის რეალიზების სიხშირე. ყველაზე უფრო დიდი სიხშირის მქონე ელემენტი განსაზღვრავს მოცემული კლასის წარმომადგენელს /ა/ სიმბოლოსათვის. ნაჩვენებ ცხრილში ყვითელი ფერით გამოყოფილია ერთ-ერთი კლასი, ხოლო წითელი ფერით მისი მაქსიმალური ელემენტი. ეს არის 1465 და იგი შეესაბამება "რამ" ტრიფონს. ანალოგიურად განისაზღვრება სხვა კლასების წარმომადგენლები /ა/-თვის და ყველა დანარჩენი სიმბოლოსათვის.

აღნიშნული გამოთვლების კუსტარულად წარმოება მოითხოვს დიდ დროს. ყველა ზემოთ აღწერილი შედეგი მიღებულია MatLab-ის გარემოში m პროგრამირების ენაზე, ჩვენს მიერ სპეციალურად ამ მიზნით შექმნილი პროგრამების გამოყენებით. პროგრამულად ხდება კლასების წარმომადგენლების არჩევაც.

მესამე ცხრილში მოცემულია კლასების წარმომადგენელთა ჩამონათვალი; იგი შეიცავს 3224 ტრიფონს.

ცხრილი 3.

| | | | | | | | | | |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| '_აა' | '_აე' | '_აი' | '_აო' | '_აუ' | '_ამ' | '_არ' | '_ას' | '_აკ' | '_ეე' |
| '_ეო' | '_ეუ' | '_ეგ' | '_ენ' | '_ეს' | '_ეკ' | '_ი_' | '_ია' | '_იე' | '_იი' |
| '_იო' | '_იუ' | '_იმ' | '_ინ' | '_ის' | '_იყ' | '_ო_' | '_ოე' | '_ოი' | '_ოო' |
| '_ობ' | '_ორ' | '_ოთ' | '_ოკ' | '_უა' | '_უე' | '_უი' | '_უო' | '_უუ' | '_უმ' |
| '_უნ' | '_უნფ' | '_უკ' | '_ბ_' | '_ბა' | '_ბე' | '_ბი' | '_ბო' | '_ბუ' | '_ბგ' |
| '_ბრ' | '_ბშ' | '_ბტ' | '_ბა' | '_ბე' | '_ბი' | '_ბო' | '_ბუ' | '_ბკ' | '_ბრ' |
| '_გს' | '_გკ' | '_და' | '_დე' | '_დი' | '_დო' | '_დუ' | '_დლ' | '_დრ' | '_დბ' |
| '_დკ' | '_ვა' | '_ვე' | '_ვი' | '_ვო' | '_ვუ' | '_ვბ' | '_ვნ' | '_ვფ' | '_ვკ' |
| '_ზა' | '_ზე' | '_ზი' | '_ზო' | '_ზუ' | '_ზლ' | '_ზრ' | '_თ_' | '_თა' | '_თე' |
| '_თი' | '_თო' | '_თუ' | '_თვ' | '_თრ' | '_თქ' | '_კა' | '_კე' | '_კი' | '_კო' |
| '_კუ' | '_კვ' | '_კლ' | '_კს' | '_ლა' | '_ლე' | '_ლი' | '_ლო' | '_ლუ' | '_ლლ' |
| '_ლს' | '_ლკ' | '_მა' | '_მე' | '_მი' | '_მო' | '_მუ' | '_მგ' | '_მდ' | '_მხ' |
| '_მკ' | '_ნა' | '_ნე' | '_ნი' | '_ნო' | '_ნუ' | '_ნგ' | '_ნდ' | '_ნთ' | '_ნტ' |
| '_პა' | '_პე' | '_პი' | '_პო' | '_პუ' | '_პრ' | '_პშ' | '_პტ' | '_ჟა' | '_ჟე' |
| '_ჟი' | '_ჟო' | '_ჟუ' | '_ჟლ' | '_ჟრ' | '_რა' | '_რე' | '_რი' | '_რო' | '_რუ' |
| '_რვ' | '_რძ' | '_რჩ' | '_რკ' | '_სა' | '_სე' | '_სი' | '_სო' | '_სუ' | '_სვ' |
| '_სრ' | '_სხ' | '_სტ' | '_ტა' | '_ტე' | '_ტი' | '_ტო' | '_ტუ' | '_ტბ' | '_ტრ' |
| '_ტფ' | '_ტყ' | '_ფა' | '_ფე' | '_ფი' | '_ფო' | '_ფუ' | '_ფრ' | '_ფს' | '_ფკ' |
| '_ქა' | '_ქე' | '_ქი' | '_ქო' | '_ქუ' | '_ქვ' | '_ქრ' | '_ქს' | '_ლა' | '_ლე' |
| '_ღი' | '_ღო' | '_ღუ' | '_ღვ' | '_ღრ' | '_ყა' | '_ყე' | '_ყი' | '_ყო' | '_ყუ' |
| '_ყვ' | '_ყრ' | '_ყშ' | '_შა' | '_შე' | '_ში' | '_შო' | '_შუ' | '_შვ' | '_შრ' |
| '_შთ' | '_შტ' | '_ჩა' | '_ჩე' | '_ჩი' | '_ჩო' | '_ჩუ' | '_ჩვ' | '_ჩრ' | '_ჩხ' |
| '_ცა' | '_ცე' | '_ცი' | '_ცო' | '_ცუ' | '_ცვ' | '_ცნ' | '_ცხ' | '_ცკ' | '_ცა' |
| '_ძე' | '_ძი' | '_ძო' | '_ძუ' | '_ძვ' | '_ძლ' | '_წა' | '_წე' | '_წი' | '_წო' |
| '_წუ' | '_წვ' | '_წლ' | '_წყ' | '_ჭა' | '_ჭე' | '_ჭი' | '_ჭო' | '_ჭუ' | '_ჭვ' |
| '_ჭრ' | '_ჭყ' | '_ხა' | '_ხე' | '_ხი' | '_ხო' | '_ხუ' | '_ხმ' | '_ხდ' | '_ხშ' |
| '_ხტ' | '_ჯა' | '_ჯე' | '_ჯი' | '_ჯო' | '_ჯუ' | '_ჯგ' | '_ჯნ' | '_ჰა' | '_ჰე' |
| '_ჰი' | '_ჰო' | '_ჰუ' | '_ჰგ' | '_ჰნ' | '_ჰქ' | '_ჰყ' | '_აა_' | '_ააა' | '_ააე' |
| '_ააი' | '_ააო' | '_ააუ' | '_აამ' | '_აარ' | '_აათ' | '_ააკ' | '_აე_' | '_აეი' | '_აეუ' |
| '_აებ' | '_აერ' | '_აემ' | '_აეკ' | '_აი_' | '_აია' | '_აიე' | '_აიი' | '_აიო' | '_აიუ' |
| '_აიმ' | '_აინ' | '_აის' | '_აიწ' | '_აო_' | '_აოა' | '_აოე' | '_აოი' | '_აოუ' | '_აობ' |
| '_აოდ' | '_აოს' | '_აოჭ' | '_აუ_' | '_აუა' | '_აუე' | '_აუო' | '_აუუ' | '_აუბ' | '_აურ' |
| '_აუხ' | '_აუკ' | '_აბ_' | '_აბა' | '_აბე' | '_აბი' | '_აბო' | '_აბუ' | '_აბმ' | '_აბრ' |
| '_აბს' | '_აბჭ' | '_აგ_' | '_აგა' | '_აგე' | '_აგი' | '_აგო' | '_აგუ' | '_აგვ' | '_აგრ' |
| '_აგს' | '_აგკ' | '_ად_' | '_ადა' | '_ადე' | '_ადი' | '_ადო' | '_ადუ' | '_ადგ' | '_ადრ' |
| '_ადს' | '_ადყ' | '_ავ_' | '_ავა' | '_ავე' | '_ავი' | '_ავო' | '_ავუ' | '_ავმ' | '_ავლ' |
| '_ავს' | '_ავკ' | '_აზ_' | '_აზა' | '_აზე' | '_აზი' | '_აზო' | '_აზუ' | '_აზლ' | '_აზრ' |
| '_აზთ' | '_აზკ' | '_ათ_' | '_ათა' | '_ათე' | '_ათი' | '_ათო' | '_ათუ' | '_ათვ' | '_ათლ' |
| '_ათშ' | '_ათწ' | '_აკ_' | '_აკა' | '_აკე' | '_აკი' | '_აკო' | '_აკუ' | '_აკვ' | '_აკრ' |
| '_აკს' | '_აკტ' | '_ალ_' | '_ალა' | '_ალე' | '_ალი' | '_ალო' | '_ალუ' | '_ალბ' | '_ალდ' |

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 'აღს' | 'აღვ' | 'ამ_' | 'ამა' | 'ამე' | 'ამი' | 'ამო' | 'ამუ' | 'ამბ' | 'ამდ' |
| 'ამხ' | 'ამწ' | 'ან_' | 'ანა' | 'ანე' | 'ანი' | 'ანო' | 'ანუ' | 'ანმ' | 'ანდ' |
| 'ანს' | 'ანტ' | 'აპ_' | 'აპა' | 'აპე' | 'აპი' | 'აპო' | 'აპუ' | 'აპვ' | 'აპრ' |
| 'აპს' | 'აპყ' | 'აჟა' | 'აჟე' | 'აჟი' | 'აჟო' | 'აჟღ' | 'აჟღ' | 'აჟს' | 'აჟკ' |
| 'არ_' | 'არა' | 'არე' | 'არი' | 'არო' | 'არუ' | 'არმ' | 'არდ' | 'ართ' | 'არტ' |
| 'ას_' | 'ასა' | 'ასე' | 'ასი' | 'ასო' | 'ასუ' | 'ასვ' | 'ასრ' | 'ასხ' | 'ასტ' |
| 'ატ_' | 'ატა' | 'ატე' | 'ატი' | 'ატო' | 'ატუ' | 'ატვ' | 'ატრ' | 'ატთ' | 'ატყ' |
| 'აფ_' | 'აფა' | 'აფე' | 'აფი' | 'აფო' | 'აფუ' | 'აფვ' | 'აფრ' | 'აფხ' | 'აფტ' |
| 'აქ_' | 'აქა' | 'აქე' | 'აქი' | 'აქო' | 'აქუ' | 'აქვ' | 'აქრ' | 'აქს' | 'აქტ' |
| 'აღ_' | 'აღა' | 'აღე' | 'აღი' | 'აღო' | 'აღუ' | 'აღმ' | 'აღნ' | 'აღს' | 'აღწ' |
| 'აყა' | 'აყე' | 'აყი' | 'აყო' | 'აყუ' | 'აყვ' | 'აყრ' | 'აყთ' | 'აშ_' | 'აშა' |
| 'აშე' | 'აში' | 'აშო' | 'აშუ' | 'აშვ' | 'აშრ' | 'აშქ' | 'აშკ' | 'აჩა' | 'აჩე' |
| 'აჩი' | 'აჩო' | 'აჩუ' | 'აჩვ' | 'აჩნ' | 'აჩქ' | 'აჩკ' | 'აც_' | 'აცა' | 'აცე' |
| 'აცი' | 'აცი' | 'აცუ' | 'აცვ' | 'აცნ' | 'აცხ' | 'აცტ' | 'ად_' | 'ადა' | 'ადე' |
| 'ადი' | 'ადო' | 'ადუ' | 'ადვ' | 'ადლ' | 'აწ_' | 'აწა' | 'აწე' | 'აწი' | 'აწო' |
| 'აწუ' | 'აწვ' | 'აწლ' | 'აწც' | 'აწყ' | 'აჭ_' | 'აჭა' | 'აჭე' | 'აჭი' | 'აჭუ' |
| 'აჭვ' | 'აჭრ' | 'აჭკ' | 'ახ_' | 'ახა' | 'ახე' | 'ახი' | 'ახო' | 'ახუ' | 'ახვ' |
| 'ახლ' | 'ახს' | 'ახტ' | 'აჯ_' | 'აჯა' | 'აჯე' | 'აჯი' | 'აჯო' | 'აჯუ' | 'აჯგ' |
| 'აჯღ' | 'აჰ_' | 'აჰა' | 'აჰი' | 'აჰუ' | 'აჰმ' | 'აჰდ' | 'აჰქ' | 'აჰკ' | 'ა_ა' |
| 'ა_ე' | 'ა_ი' | 'ა_ო' | 'ა_უ' | 'ა_მ' | 'ა_ღ' | 'ა_ს' | 'ა_კ' | 'ეა_' | 'ეაა' |
| 'ეაო' | 'ეამ' | 'ეან' | 'ეას' | 'ეატ' | 'ეე_' | 'ეეე' | 'ეეი' | 'ეებ' | 'ეერ' |
| 'ეეს' | 'ეეტ' | 'ეი_' | 'ეია' | 'ეიე' | 'ეიძ' | 'ეიძ' | 'ეის' | 'ეიტ' | 'ეო_' |
| 'ეოა' | 'ეოე' | 'ეობ' | 'ეორ' | 'ეოს' | 'ეოტ' | 'ეუ_' | 'ეუე' | 'ეუმ' | 'ეულ' |
| 'ეუს' | 'ეუწ' | 'ებ_' | 'ება' | 'ებე' | 'ები' | 'ებო' | 'ებუ' | 'ებმ' | 'ებდ' |
| 'ებს' | 'ებწ' | 'ებ_' | 'ება' | 'ებე' | 'ები' | 'ებო' | 'ებუ' | 'ებვ' | 'ებრ' |
| 'ებს' | 'ებწ' | 'ეგ_' | 'ეგა' | 'ეგე' | 'ეგი' | 'ეგო' | 'ეგუ' | 'ეგვ' | 'ეგრ' |
| 'ეგს' | 'ეგჭ' | 'ედ_' | 'ედა' | 'ედე' | 'ედი' | 'ედო' | 'ედუ' | 'ედვ' | 'ედრ' |
| 'ედს' | 'ედწ' | 'ევ_' | 'ევა' | 'ევე' | 'ევი' | 'ევო' | 'ევუ' | 'ევმ' | 'ევრ' |
| 'ევს' | 'ევყ' | 'ეზ_' | 'ეზა' | 'ეზე' | 'ეზი' | 'ეზო' | 'ეზუ' | 'ეზლ' | 'ეზრ' |
| 'ეზს' | 'ეთ_' | 'ეთა' | 'ეთე' | 'ეთი' | 'ეთო' | 'ეთუ' | 'ეთვ' | 'ეთნ' | 'ეთშ' |
| 'ეთკ' | 'ეკ_' | 'ეკა' | 'ეკე' | 'ეკი' | 'ეკო' | 'ეკუ' | 'ეკვ' | 'ეკლ' | 'ეკს' |
| 'ეკვ' | 'ელ_' | 'ელა' | 'ელე' | 'ელი' | 'ელო' | 'ელუ' | 'ელმ' | 'ელნ' | 'ელს' |
| 'ელწ' | 'ემ_' | 'ემა' | 'ემე' | 'ემი' | 'ემო' | 'ემუ' | 'ემბ' | 'ემდ' | 'ემთ' |
| 'ემკ' | 'ენ_' | 'ენა' | 'ენე' | 'ენი' | 'ენო' | 'ენუ' | 'ენგ' | 'ენდ' | 'ენს' |
| 'ენტ' | 'ეპ_' | 'ეპა' | 'ეპე' | 'ეპი' | 'ეპო' | 'ეპუ' | 'ეპრ' | 'ეპს' | 'ეპტ' |
| 'ეჟა' | 'ეჟი' | 'ერ_' | 'ერა' | 'ერე' | 'ერი' | 'ერო' | 'ერუ' | 'ერმ' | 'ერძ' |
| 'ერთ' | 'ერკ' | 'ეს_' | 'ესა' | 'ესე' | 'ესი' | 'ესო' | 'ესუ' | 'ესმ' | 'ესრ' |
| 'ესხ' | 'ესტ' | 'ეტ_' | 'ეტა' | 'ეტე' | 'ეტი' | 'ეტო' | 'ეტუ' | 'ეტვ' | 'ეტრ' |
| 'ეტშ' | 'ეტყ' | 'ეფ_' | 'ეფა' | 'ეფე' | 'ეფი' | 'ეფო' | 'ეფუ' | 'ეფრ' | 'ეფხ' |
| 'ექ_' | 'ექა' | 'ექე' | 'ექი' | 'ექო' | 'ექუ' | 'ექმ' | 'ექნ' | 'ექს' | 'ექტ' |
| 'ელ_' | 'ელა' | 'ელე' | 'ელი' | 'ელო' | 'ელუ' | 'ელვ' | 'ელრ' | 'ელწ' | 'ეყ_' |
| 'ეყა' | 'ეყე' | 'ეყი' | 'ეყო' | 'ეყუ' | 'ეყვ' | 'ემ_' | 'ემა' | 'ემე' | 'ემი' |
| 'ემო' | 'ემუ' | 'ემვ' | 'ემლ' | 'ემს' | 'ემტ' | 'ეჩა' | 'ეჩე' | 'ეჩი' | 'ეჩო' |
| 'ეჩვ' | 'ეჩნ' | 'ეჩხ' | 'ეც_' | 'ეცა' | 'ეცე' | 'ეცი' | 'ეცო' | 'ეცუ' | 'ეცვ' |
| 'ეცნ' | 'ეცხ' | 'ეძა' | 'ეძე' | 'ეძი' | 'ეძუ' | 'ეძგ' | 'ეძლ' | 'ეწ_' | 'ეწა' |

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 'ენე' | 'ენი' | 'ენო' | 'ენუ' | 'ენვ' | 'ენრ' | 'ენს' | 'ენყ' | 'ექ_' | 'ექა' |
| 'ექე' | 'ექი' | 'ექო' | 'ექუ' | 'ექვ' | 'ექდ' | 'ექს' | 'ექყ' | 'ეხ_' | 'ეხა' |
| 'ეხი' | 'ეხო' | 'ეხუ' | 'ეხვ' | 'ეხზ' | 'ეხბ' | 'ეხს' | 'ეჯა' | 'ეჯი' | 'ეჯო' |
| 'ეჯბ' | 'ეჯდ' | 'ეჯს' | 'ეკ_' | 'ეკე' | 'ეკი' | 'ეკრ' | 'ეკყ' | 'ეა_' | 'ეაე' |
| 'ე_ო' | 'ე_ო' | 'ე_უ' | 'ე_მ' | 'ე_დ' | 'ე_ს' | 'ე_კ' | 'ია_' | 'იაა' | 'იაე' |
| 'იაი' | 'იაო' | 'იაუ' | 'იამ' | 'იან' | 'იათ' | 'იაკ' | 'იე_' | 'იეა' | 'იეი' |
| 'იებ' | 'იერ' | 'იექ' | 'იეტ' | 'იი_' | 'იია' | 'იიო' | 'იიდ' | 'იიდ' | 'იის' |
| 'იიწ' | 'იო_' | 'იოა' | 'იოე' | 'იოვ' | 'იონ' | 'იოს' | 'იოტ' | 'იუ_' | 'იუა' |
| 'იუე' | 'იუი' | 'იუმ' | 'იურ' | 'იუს' | 'იუტ' | 'იბ_' | 'იბა' | 'იბე' | 'იბი' |
| 'იბო' | 'იბუ' | 'იბვ' | 'იბლ' | 'იბს' | 'იბჭ' | 'იგ_' | 'იგა' | 'იგე' | 'იგი' |
| 'იგო' | 'იგუ' | 'იგვ' | 'იგნ' | 'იგშ' | 'იდ_' | 'იდა' | 'იდე' | 'იდი' | 'იდო' |
| 'იდუ' | 'იდგ' | 'იდრ' | 'იდა' | 'იდტ' | 'ივ_' | 'ივა' | 'ივე' | 'ივი' | 'ივო' |
| 'ივუ' | 'ივმ' | 'ივრ' | 'ივს' | 'ივტ' | 'იზ_' | 'იზა' | 'იზე' | 'იზი' | 'იზო' |
| 'იზუ' | 'იზმ' | 'იზნ' | 'იზს' | 'ით_' | 'ითა' | 'ითე' | 'ითი' | 'ითო' | 'ითუ' |
| 'ითვ' | 'ითრ' | 'ითხ' | 'ითკ' | 'იკ_' | 'იკა' | 'იკე' | 'იკი' | 'იკო' | 'იკუ' |
| 'იკვ' | 'იკლ' | 'იკს' | 'იკტ' | 'ილ_' | 'ილა' | 'ილე' | 'ილი' | 'ილო' | 'ილუ' |
| 'ილმ' | 'ილზ' | 'ილს' | 'ილტ' | 'იმ_' | 'იმა' | 'იმე' | 'იმი' | 'იმო' | 'იმუ' |
| 'იმბ' | 'იმდ' | 'იმთ' | 'იმყ' | 'ინ_' | 'ინა' | 'ინე' | 'ინი' | 'ინო' | 'ინუ' |
| 'ინგ' | 'ინდ' | 'ინც' | 'ინტ' | 'იპ_' | 'იპა' | 'იპე' | 'იპი' | 'იპო' | 'იპუ' |
| 'იპლ' | 'იპს' | 'იპყ' | 'იჟ_' | 'იჟა' | 'იჟე' | 'იჟი' | 'იჟო' | 'იჟუ' | 'იჟდ' |
| 'იჟრ' | 'ირ_' | 'ირა' | 'ირე' | 'ირი' | 'ირო' | 'ირუ' | 'ირვ' | 'ირდ' | 'ირს' |
| 'ირკ' | 'ის_' | 'ისა' | 'ისე' | 'ისი' | 'ისო' | 'ისუ' | 'ისმ' | 'ისრ' | 'ისთ' |
| 'ისტ' | 'იტ_' | 'იტა' | 'იტე' | 'იტი' | 'იტო' | 'იტუ' | 'იტვ' | 'იტნ' | 'იტს' |
| 'იტყ' | 'იფ_' | 'იფა' | 'იფე' | 'იფი' | 'იფო' | 'იფუ' | 'იფრ' | 'იფშ' | 'იფტ' |
| 'იქ_' | 'იქა' | 'იქე' | 'იქი' | 'იქო' | 'იქუ' | 'იქმ' | 'იქრ' | 'იქს' | 'იქტ' |
| 'ილა' | 'ილე' | 'ილი' | 'ილო' | 'ილუ' | 'იღვ' | 'იღრ' | 'იღთ' | 'იღწ' | 'იყ_' |
| 'იყა' | 'იყე' | 'იყი' | 'იყო' | 'იყუ' | 'იყვ' | 'იყრ' | 'იყს' | 'იყტ' | 'იმ_' |
| 'იშა' | 'იშე' | 'იში' | 'იშო' | 'იშუ' | 'იშვ' | 'იშნ' | 'იშხ' | 'იშკ' | 'იჩ_' |
| 'იჩა' | 'იჩე' | 'იჩი' | 'იჩო' | 'იჩუ' | 'იჩვ' | 'იჩნ' | 'იჩქ' | 'იჩკ' | 'იც_' |
| 'იცა' | 'იცე' | 'იცი' | 'იცო' | 'იცუ' | 'იცვ' | 'იცნ' | 'იცხ' | 'იცკ' | 'იდ_' |
| 'იძა' | 'იძე' | 'იძი' | 'იძო' | 'იძუ' | 'იძგ' | 'იძლ' | 'იწ_' | 'იწა' | 'იწე' |
| 'იწი' | 'იწო' | 'იწუ' | 'იწვ' | 'იწრ' | 'იწს' | 'იწყ' | 'იჭ_' | 'იჭა' | 'იჭე' |
| 'იჭი' | 'იჭო' | 'იჭუ' | 'იჭვ' | 'იჭრ' | 'იჭყ' | 'იხ_' | 'იხა' | 'იხე' | 'იხი' |
| 'იხო' | 'იხუ' | 'იხვ' | 'იხდ' | 'იხს' | 'იხტ' | 'იჯ_' | 'იჯა' | 'იჯე' | 'იჯი' |
| 'იჯო' | 'იჯვ' | 'იჯნ' | 'იჯს' | 'იპა' | 'იპქ' | 'იპყ' | 'ი_ა' | 'ი_ე' | 'ი_ი' |
| 'ი_ო' | 'ი_უ' | 'ი_მ' | 'ი_დ' | 'ი_ს' | 'ი_კ' | 'ოა_' | 'ოად' | 'ოად' | 'ოაქ' |
| 'ოაყ' | 'ოე_' | 'ოეა' | 'ოეი' | 'ოეუ' | 'ოებ' | 'ოელ' | 'ოეთ' | 'ოეტ' | 'ოი_' |
| 'ოია' | 'ოიე' | 'ოიო' | 'ოიმ' | 'ოინ' | 'ოივ' | 'ოიყ' | 'ოო_' | 'ოოა' | 'ოობ' |
| 'ოოდ' | 'ოოფ' | 'ოუ_' | 'ოუა' | 'ოუე' | 'ოუვ' | 'ოურ' | 'ოუს' | 'ოუკ' | 'ობ_' |
| 'ობა' | 'ობე' | 'ობი' | 'ობო' | 'ობუ' | 'ობმ' | 'ობრ' | 'ობს' | 'ობ_' | 'ობა' |
| 'ობე' | 'ობი' | 'ობო' | 'ობუ' | 'ობვ' | 'ობრ' | 'ობთ' | 'ობტ' | 'ოდ_' | 'ოდა' |
| 'ოდე' | 'ოდი' | 'ოდო' | 'ოდუ' | 'ოდგ' | 'ოდნ' | 'ოდშ' | 'ოვ_' | 'ოვა' | 'ოვე' |
| 'ოვი' | 'ოვო' | 'ოვუ' | 'ოვმ' | 'ოვრ' | 'ოვს' | 'ოვკ' | 'ოზ_' | 'ოზა' | 'ოზე' |
| 'ოზი' | 'ოზო' | 'ოზუ' | 'ოზდ' | 'ოზდ' | 'ოზს' | 'ოთ_' | 'ოთა' | 'ოთე' | 'ოთი' |

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 'ოთო' | 'ოთუ' | 'ოთმ' | 'ოთლ' | 'ოთხ' | 'ოთკ' | 'ოკ_' | 'ოკა' | 'ოკე' | 'ოკი' |
| 'ოკო' | 'ოკუ' | 'ოკვ' | 'ოკლ' | 'ოკმ' | 'ოლ_' | 'ოლა' | 'ოლე' | 'ოლი' | 'ოლო' |
| 'ოლუ' | 'ოლმ' | 'ოლნ' | 'ოლხ' | 'ოლკ' | 'ომ_' | 'ომა' | 'ომე' | 'ომი' | 'ომო' |
| 'ომუ' | 'ომმ' | 'ომლ' | 'ომთ' | 'ომპ' | 'ონ_' | 'ონა' | 'ონე' | 'ონი' | 'ონო' |
| 'ონუ' | 'ონვ' | 'ონდ' | 'ონს' | 'ონტ' | 'ოპ_' | 'ოპა' | 'ოპე' | 'ოპი' | 'ოპო' |
| 'ოპუ' | 'ოპმ' | 'ოპრ' | 'ოპს' | 'ოპტ' | 'ოჟ_' | 'ოჟა' | 'ოჟე' | 'ოჟი' | 'ოჟო' |
| 'ოჟრ' | 'ორ_' | 'ორა' | 'ორე' | 'ორი' | 'ორო' | 'ორუ' | 'ორმ' | 'ორდ' | 'ორც' |
| 'ორტ' | 'ოს_' | 'ოსა' | 'ოსე' | 'ოსი' | 'ოსო' | 'ოსუ' | 'ოსვ' | 'ოსნ' | 'ოსხ' |
| 'ოსტ' | 'ოტ_' | 'ოტა' | 'ოტე' | 'ოტი' | 'ოტო' | 'ოტუ' | 'ოტბ' | 'ოტლ' | 'ოტს' |
| 'ოტყ' | 'ოფ_' | 'ოფა' | 'ოფე' | 'ოფი' | 'ოფო' | 'ოფუ' | 'ოფვ' | 'ოფლ' | 'ოფს' |
| 'ოქ_' | 'ოქა' | 'ოქე' | 'ოქი' | 'ოქო' | 'ოქუ' | 'ოქმ' | 'ოქრ' | 'ოქც' | 'ოქტ' |
| 'ოდ_' | 'ოდა' | 'ოდე' | 'ოდი' | 'ოდო' | 'ოდუ' | 'ოდვ' | 'ოდლ' | 'ოდწ' | 'ოყ_' |
| 'ოყა' | 'ოყე' | 'ოყი' | 'ოყო' | 'ოყუ' | 'ოყვ' | 'ოყრ' | 'ომ_' | 'ომა' | 'ომე' |
| 'ომი' | 'ომო' | 'ომუ' | 'ომვ' | 'ომლ' | 'ომხ' | 'ომკ' | 'ოჩ_' | 'ოჩა' | 'ოჩე' |
| 'ოჩი' | 'ოჩო' | 'ოჩნ' | 'ოჩხ' | 'ოც_' | 'ოცა' | 'ოცე' | 'ოცი' | 'ოცო' | 'ოცუ' |
| 'ოცვ' | 'ოცდ' | 'ოცხ' | 'ოცტ' | 'ომ_' | 'ომა' | 'ომე' | 'ომი' | 'ომო' | 'ომუ' |
| 'ომლ' | 'ომრ' | 'ონ_' | 'ონა' | 'ონე' | 'ონი' | 'ონო' | 'ონმ' | 'ონწ' | 'ონყ' |
| 'ოჭ_' | 'ოჭა' | 'ოჭე' | 'ოჭი' | 'ოჭო' | 'ოჭუ' | 'ოჭრ' | 'ოჭს' | 'ოჭკ' | 'ობ_' |
| 'ობა' | 'ობე' | 'ობი' | 'ობო' | 'ობუ' | 'ობვ' | 'ობდ' | 'ობს' | 'ობტ' | 'ოჯ_' |
| 'ოჯა' | 'ოჯე' | 'ოჯი' | 'ოჯო' | 'ოჯდ' | 'ოპ_' | 'ოპა' | 'ოპო' | 'ოპგ' | 'ოპქ' |
| 'ოპყ' | 'ო_ა' | 'ო_ე' | 'ო_ი' | 'ო_ო' | 'ო_უ' | 'ო_მ' | 'ო_დ' | 'ო_ს' | 'ო_კ' |
| 'უა_' | 'უაა' | 'უაი' | 'უაუ' | 'უამ' | 'უალ' | 'უახ' | 'უაწ' | 'უე_' | 'უებ' |
| 'უელ' | 'უეც' | 'უექ' | 'უი_' | 'უია' | 'უიი' | 'უიო' | 'უიმ' | 'უილ' | 'უით' |
| 'უიტ' | 'უო_' | 'უოა' | 'უოო' | 'უოუ' | 'უოდ' | 'უოს' | 'უოკ' | 'უუ_' | 'უუმ' |
| 'უულ' | 'უუს' | 'უბ_' | 'უბა' | 'უბე' | 'უბი' | 'უბო' | 'უბუ' | 'უბრ' | 'უბს' |
| 'უგა' | 'უგე' | 'უგი' | 'უგო' | 'უგუ' | 'უგვ' | 'უგდ' | 'უდ_' | 'უდა' | 'უდე' |
| 'უდი' | 'უდო' | 'უდუ' | 'უდგ' | 'უდრ' | 'უდს' | 'უვა' | 'უვე' | 'უვი' | 'უვო' |
| 'უვლ' | 'უზა' | 'უზე' | 'უზი' | 'უზო' | 'უზუ' | 'უზმ' | 'უზრ' | 'უზშ' | 'უთ_' |
| 'უთა' | 'უთე' | 'უთი' | 'უთო' | 'უთუ' | 'უთვ' | 'უთრ' | 'უთხ' | 'უთტ' | 'უკ_' |
| 'უკა' | 'უკე' | 'უკი' | 'უკო' | 'უკუ' | 'უკვ' | 'უკრ' | 'უკმ' | 'ულ_' | 'ულა' |
| 'ულე' | 'ული' | 'ულო' | 'ულუ' | 'ულმ' | 'ულდ' | 'ულშ' | 'ულტ' | 'უმ_' | 'უმა' |
| 'უმე' | 'უმი' | 'უმო' | 'უმუ' | 'უმბ' | 'უმლ' | 'უმც' | 'უმწ' | 'უნ_' | 'უნა' |
| 'უნე' | 'უნი' | 'უნო' | 'უნუ' | 'უნვ' | 'უნდ' | 'უნთ' | 'უნტ' | 'უპ_' | 'უპა' |
| 'უპე' | 'უპი' | 'უპო' | 'უპუ' | 'უპვ' | 'უპრ' | 'უპს' | 'უპტ' | 'უჟე' | 'ურ_' |
| 'ურა' | 'ურე' | 'ური' | 'ურო' | 'ურუ' | 'ურგ' | 'ურდ' | 'ურქ' | 'ურტ' | 'უს_' |
| 'უსა' | 'უსე' | 'უსი' | 'უსო' | 'უსუ' | 'უსმ' | 'უსზ' | 'უსხ' | 'უსტ' | 'უტ_' |
| 'უტა' | 'უტე' | 'უტი' | 'უტო' | 'უტუ' | 'უტმ' | 'უტზ' | 'უტს' | 'უტკ' | 'უფ_' |
| 'უფა' | 'უფე' | 'უფი' | 'უფო' | 'უფუ' | 'უფმ' | 'უფრ' | 'უფს' | 'უფტ' | 'უქა' |
| 'უქე' | 'უქი' | 'უქო' | 'უქუ' | 'უქმ' | 'უქრ' | 'უქც' | 'უქტ' | 'უღ_' | 'უღა' |
| 'უღე' | 'უღი' | 'უღო' | 'უღუ' | 'უღვ' | 'უღლ' | 'უღს' | 'უღწ' | 'უყა' | 'უყე' |
| 'უყი' | 'უყო' | 'უყუ' | 'უყვ' | 'უყრ' | 'უმ_' | 'უმა' | 'უმე' | 'უმი' | 'უმო' |
| 'უმუ' | 'უმვ' | 'უმლ' | 'უმქ' | 'უმტ' | 'უჩ_' | 'უჩა' | 'უჩე' | 'უჩი' | 'უჩო' |
| 'უჩუ' | 'უჩვ' | 'უჩნ' | 'უჩქ' | 'უჩკ' | 'უც_' | 'უცა' | 'უცე' | 'უცი' | 'უცუ' |
| 'უცვ' | 'უცნ' | 'უცხ' | 'უძ_' | 'უძა' | 'უძე' | 'უძი' | 'უძუ' | 'უძვ' | 'უძლ' |

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 'უწ_' | 'უწე' | 'უწი' | 'უწო' | 'უწუ' | 'უწვ' | 'უწნ' | 'უწყ' | 'უჭ_' | 'უჭა' |
| 'უჭე' | 'უჭი' | 'უჭო' | 'უჭუ' | 'უჭრ' | 'უჭშ' | 'უხ_' | 'უხა' | 'უხე' | 'უხი' |
| 'უხო' | 'უხუ' | 'უხვ' | 'უხლ' | 'უხჩ' | 'უხტ' | 'უჯ_' | 'უჯა' | 'უჯე' | 'უჯი' |
| 'უჯო' | 'უჯუ' | 'უჯრ' | 'უჯს' | 'უჰ_' | 'უჰა' | 'უ_ა' | 'უ_ე' | 'უ_ი' | 'უ_ო' |
| 'უ_უ' | 'უ_მ' | 'უ_რ' | 'უ_ს' | 'უ_კ' | 'მა_' | 'ნაა' | 'მაე' | 'მაი' | 'რაო' |
| 'ნაუ' | 'რამ' | 'მარ' | 'მას' | 'რატ' | 'მე_' | 'რეა' | 'ნეე' | 'ნეი' | 'მეო' |
| 'ლეუ' | 'ლებ' | 'მელ' | 'ლექ' | 'მეტ' | 'ლი_' | 'ლია' | 'ნიე' | 'რიი' | 'რიო' |
| 'რიუ' | 'რივ' | 'ნილ' | 'ნის' | 'ლიტ' | 'რო_' | 'მოა' | 'ლოე' | 'მოი' | 'ლოო' |
| 'მოუ' | 'რომ' | 'ლოდ' | 'მოს' | 'მოყ' | 'ნუ_' | 'რუა' | 'რუე' | 'ლუი' | 'მუო' |
| 'ნუუ' | 'რუმ' | 'ლურ' | 'ლუს' | 'რუტ' | 'მბ_' | 'რბა' | 'მბე' | 'რბი' | 'მბო' |
| 'მბუ' | 'მბმ' | 'რბლ' | 'რბთ' | 'ნგ_' | 'რგა' | 'რგე' | 'რგი' | 'რგო' | 'რგუ' |
| 'რგმ' | 'ნგლ' | 'რგს' | 'ნგტ' | 'რდ_' | 'ნდა' | 'მდე' | 'რდი' | 'ნდო' | 'რდუ' |
| 'მდგ' | 'ნდრ' | 'ნდს' | 'რდტ' | 'რვა' | 'რვე' | 'ნვი' | 'ლვო' | 'რვუ' | 'ნვმ' |
| 'ნვლ' | 'ნვს' | 'რვე' | 'ლზ_' | 'მზა' | 'ლზე' | 'მზი' | 'ნზო' | 'ნზუ' | 'ლზმ' |
| 'ნზრ' | 'რზს' | 'რთ_' | 'ლთა' | 'რთე' | 'რთი' | 'რთო' | 'რთუ' | 'რთვ' | 'რთლ' |
| 'მთხ' | 'ნთკ' | 'რკ_' | 'რკა' | 'ლკე' | 'რკი' | 'ნკო' | 'მკუ' | 'მკვ' | 'ნკრ' |
| 'რკს' | 'ნკტ' | 'რლ_' | 'რლა' | 'მლე' | 'მლი' | 'რლო' | 'მლუ' | 'რლგ' | 'რლს' |
| 'მმ_' | 'რმა' | 'რმე' | 'რმი' | 'რმო' | 'რმუ' | 'ნმმ' | 'ლმმ' | 'ლმხ' | 'ლმწ' |
| 'რნ_' | 'მნა' | 'მნე' | 'მნი' | 'მნო' | 'მნუ' | 'ლნმ' | 'ლნზ' | 'ლნს' | 'ლნტ' |
| 'მპა' | 'მპე' | 'მპი' | 'მპო' | 'რპუ' | 'მპლ' | 'მპს' | 'მპყ' | 'რჟ_' | 'მჟა' |
| 'ნჟი' | 'მჟო' | 'რჟუ' | 'მჟლ' | 'მრ_' | 'მრა' | 'მრე' | 'მრი' | 'მრო' | 'მრუ' |
| 'მრგ' | 'მრჯ' | 'მრჩ' | 'ლს_' | 'ლსა' | 'რსე' | 'რსი' | 'მსო' | 'რსუ' | 'მსგ' |
| 'მსჯ' | 'ნსხ' | 'ნსტ' | 'რტ_' | 'ნტა' | 'ნტე' | 'ნტი' | 'ნტო' | 'ლტუ' | 'მტვ' |
| 'ნტრ' | 'ნტთ' | 'მტკ' | 'ლფ_' | 'რფა' | 'რფე' | 'რფი' | 'რფო' | 'ნფუ' | 'ლფგ' |
| 'მფლ' | 'ლფს' | 'რქ_' | 'რქა' | 'რქე' | 'რქი' | 'მქო' | 'რქუ' | 'ლქვ' | 'ნქლ' |
| 'ნქც' | 'ნქტ' | 'ლღ_' | 'რღა' | 'ლღე' | 'რღი' | 'ლღო' | 'რღუ' | 'რღვ' | 'რღნ' |
| 'მყა' | 'რყე' | 'მყი' | 'მყო' | 'მყუ' | 'მყვ' | 'მყრ' | 'რყწ' | 'რშ_' | 'მშა' |
| 'რშე' | 'ლში' | 'მშო' | 'მშუ' | 'მშვ' | 'მშრ' | 'მშფ' | 'ნშტ' | 'რჩ_' | 'რჩა' |
| 'რჩე' | 'რჩი' | 'რჩო' | 'რჩუ' | 'რჩვ' | 'მჩნ' | 'ნჩხ' | 'რც_' | 'მცა' | 'რცე' |
| 'ნცი' | 'მცო' | 'მცუ' | 'რცვ' | 'რცლ' | 'მცხ' | 'რძ_' | 'რძა' | 'რძე' | 'ნძი' |
| 'რძო' | 'ნძუ' | 'მძლ' | 'რძნ' | 'ნძთ' | 'ნწ_' | 'მწა' | 'მწე' | 'მწი' | 'ლწო' |
| 'მწუ' | 'რწმ' | 'ლწლ' | 'ნწშ' | 'რწყ' | 'ნჭა' | 'მჭე' | 'მჭი' | 'მჭო' | 'რჭუ' |
| 'მჭვ' | 'მჭრ' | 'ლზ_' | 'მზა' | 'ლზე' | 'ლზი' | 'მზო' | 'ლზუ' | 'ნზმ' | 'მზრ' |
| 'ლზთ' | 'მზტ' | 'რჯ_' | 'ნჯა' | 'რჯე' | 'რჯი' | 'ნჯო' | 'ნჯუ' | 'რჯვ' | 'რჯნ' |
| 'რჯს' | 'ნჰა' | 'მჰე' | 'რჰუ' | 'მჰლ' | 'რჰყ' | 'ნ_ა' | 'რ_ე' | 'ნ_ი' | 'მ_ო' |
| 'ნ_უ' | 'ნ_მ' | 'ნ_დ' | 'მ_ს' | 'რ_კ' | 'და_' | 'დაა' | 'დაე' | 'დაი' | 'დაო' |
| 'დაუ' | 'დამ' | 'დან' | 'დას' | 'დაპ' | 'ზე_' | 'დეა' | 'დეე' | 'ზეი' | 'დეო' |
| 'ზეუ' | 'დებ' | 'დენ' | 'დეს' | 'ზევ' | 'დი_' | 'დია' | 'ძიე' | 'ზიი' | 'დიო' |
| 'ზიუ' | 'დიგ' | 'დიდ' | 'დით' | 'ზიკ' | 'დო_' | 'დოა' | 'დოე' | 'ზოო' | 'დოუ' |
| 'ზოგ' | 'ძოდ' | 'დოთ' | 'დოკ' | 'დუ_' | 'დუა' | 'დუე' | 'დუი' | 'დუმ' | 'ზურ' |
| 'ზუს' | 'დუკ' | 'დბა' | 'ზბე' | 'დბო' | 'დბუ' | 'ზგ_' | 'დგა' | 'დგე' | 'დგი' |
| 'დგო' | 'ჯგუ' | 'დგმ' | 'დგნ' | 'ზგთ' | 'დდ_' | 'ჯდა' | 'ზდე' | 'დდი' | 'ზდო' |
| 'ზდრ' | 'დვ_' | 'დვა' | 'ძვე' | 'დვი' | 'ზვლ' | 'ძვლ' | 'დვც' | 'დზე' | 'დზო' |
| 'დზრ' | 'დთ_' | 'ზთა' | 'ძთე' | 'ძთი' | 'ჟკა' | 'ზკო' | 'დკვ' | 'ძლა' | 'ძლე' |

| | | | | | | | | | |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 'ძლი' | 'ძლო' | 'ძლუ' | 'ძლმ' | 'ძლდ' | 'ძმა' | 'ძმე' | 'ძმი' | 'ძმო' | 'ძმუ' |
| 'ზმნ' | 'ზმს' | 'ძმყ' | 'ძნ_' | 'ჯნა' | 'ძნე' | 'ძნი' | 'ძნო' | 'ძნუ' | 'ძნვ' |
| 'ძნშ' | 'ძპე' | 'ძპი' | 'ძრ_' | 'ძრა' | 'ძრე' | 'ზრი' | 'ძრო' | 'ძრუ' | 'ზრმ' |
| 'ზრდ' | 'ზრს' | 'ძრწ' | 'ძს_' | 'ძსა' | 'ძსე' | 'ძსი' | 'ძსო' | 'ძსუ' | 'ძსბ' |
| 'ძსთ' | 'ძსტ' | 'ძტ_' | 'ძტი' | 'ძტმ' | 'ძტს' | 'ძფა' | 'ძფე' | 'ძფრ' | 'ძქა' |
| 'ძქე' | 'ზქო' | 'ძქც' | 'ძღ_' | 'ზღა' | 'ძღე' | 'ძღი' | 'ძღო' | 'ზღუ' | 'ზღვ' |
| 'ძღლ' | 'ძღს' | 'ძყო' | 'ძშა' | 'ძში' | 'ძშვ' | 'ძც_' | 'ძცი' | 'ძძა' | 'ძძმ' |
| 'ძწა' | 'ძწო' | 'ზხ_' | 'ძხა' | 'ძხმ' | 'ღ_ა' | 'ღ_ე' | 'ღ_ი' | 'ღ_ო' | 'ღ_უ' |
| 'ღ_მ' | 'ღ_დ' | 'ღ_ს' | 'ღ_კ' | 'ზა_' | 'გაა' | 'გაე' | 'გაი' | 'გაო' | 'გაუ' |
| 'გამ' | 'გან' | 'ზას' | 'ვაწ' | 'ვე_' | 'ვეა' | 'ვეე' | 'ვეი' | 'გეო' | 'ვეუ' |
| 'ვებ' | 'ველ' | 'ვეშ' | 'ვეყ' | 'ზი_' | 'ვია' | 'ზიე' | 'გიი' | 'გიო' | 'გიუ' |
| 'ვიმ' | 'ვილ' | 'ზის' | 'გიკ' | 'ზო_' | 'ზოა' | 'ზოე' | 'გოი' | 'ვოო' | 'ზოუ' |
| 'ზობ' | 'ზოდ' | 'ზოს' | 'ზოტ' | 'ზუ_' | 'ვუა' | 'გუე' | 'ლუო' | 'გუუ' | 'გუმ' |
| 'ზულ' | 'გუფ' | 'ვუპ' | 'გზა' | 'ვბე' | 'გბი' | 'ვბო' | 'ლბუ' | 'ვბლ' | 'ვბრ' |
| 'ვგა' | 'ზგე' | 'ღგი' | 'ვგო' | 'ვგუ' | 'ღგვ' | 'ვგრ' | 'გგც' | 'ვღ_' | 'ზღა' |
| 'ვღე' | 'ვღი' | 'ვღო' | 'ვღუ' | 'ლღვ' | 'ზღნ' | 'ვღს' | 'ღვა' | 'გვე' | 'ღვი' |
| 'გვო' | 'გვუ' | 'გვმ' | 'ღვრ' | 'ღვთ' | 'გვჭ' | 'გზა' | 'ზზე' | 'გზი' | 'ვზო' |
| 'გზუ' | 'ვზვ' | 'გზნ' | 'ვზს' | 'ვზკ' | 'ზთ_' | 'ზთა' | 'ვთე' | 'ვთი' | 'ვთო' |
| 'ვთუ' | 'ვთვ' | 'ვთქ' | 'ვკა' | 'ვკე' | 'ვკი' | 'ვკო' | 'ვკვ' | 'ვკრ' | 'ზლ_' |
| 'ვლა' | 'ვლე' | 'ვლი' | 'ვლო' | 'ვლუ' | 'ზღვ' | 'ვლდ' | 'გლთ' | 'ზმა' | 'ღმე' |
| 'გმი' | 'ღმო' | 'ზმუ' | 'ვმმ' | 'გმნ' | 'ვმხ' | 'ღმწ' | 'გნ_' | 'ვნა' | 'ვნე' |
| 'ვნი' | 'გნო' | 'ვნუ' | 'ზნმ' | 'ვნდ' | 'გნშ' | 'ვპა' | 'ვპი' | 'ვპო' | 'ზჟა' |
| 'ზჟო' | 'ვჟლ' | 'ვრ_' | 'გრა' | 'ვრე' | 'ზრი' | 'ვრო' | 'ზრუ' | 'ღრმ' | 'გრძ' |
| 'ვრც' | 'ზრკ' | 'ზს_' | 'ზსა' | 'ვსე' | 'ვსი' | 'ვსო' | 'ვსუ' | 'ვსმ' | 'გსლ' |
| 'ვსხ' | 'ვსტ' | 'ვტ_' | 'ღტა' | 'ვტე' | 'ვტი' | 'ვტო' | 'გტმ' | 'ვტრ' | 'გტს' |
| 'ვტყ' | 'ვფა' | 'ვფი' | 'ვფო' | 'ვფუ' | 'ვფრ' | 'ვქა' | 'ვქე' | 'ვქი' | 'ვქო' |
| 'ღქმ' | 'ვქნ' | 'ვქს' | 'ზღა' | 'ზღე' | 'ზღუ' | 'ვღვ' | 'ზღდ' | 'ვყა' | 'ვყე' |
| 'ვყი' | 'ვყო' | 'ვყუ' | 'ვყვ' | 'ვყლ' | 'ვშ_' | 'ვშა' | 'ვშე' | 'ზში' | 'ვშო' |
| 'ვშუ' | 'ვშვ' | 'ვშნ' | 'ღშფ' | 'ვშტ' | 'ვჩა' | 'ვჩე' | 'ზჩუ' | 'ვჩნ' | 'ვცა' |
| 'ვცე' | 'ვცი' | 'ვცო' | 'ვცუ' | 'ვცვ' | 'ვცდ' | 'ვცქ' | 'ვძა' | 'ვძე' | 'ვძი' |
| 'ვძვ' | 'ვძლ' | 'ზწ_' | 'გწა' | 'ღწე' | 'ღწი' | 'ვწო' | 'ვწვ' | 'ვწყ' | 'ვჭა' |
| 'ვჭე' | 'ვჭი' | 'ზჭო' | 'ღჭუ' | 'ვჭრ' | 'ვჭყ' | 'ვზა' | 'ვზე' | 'ვზო' | 'ვზუ' |
| 'ვზვ' | 'ვზდ' | 'ვზს' | 'ვზტ' | 'გჯ_' | 'ვჯა' | 'გჯე' | 'ზჯი' | 'ზჯლ' | 'ვჯდ' |
| 'ვჯს' | 'ვჯკ' | 'ვ_ა' | 'ვ_ე' | 'ზ_ი' | 'გ_ო' | 'ვ_უ' | 'ვ_მ' | 'ვ_ჯ' | 'ვ_ს' |
| 'ვ_კ' | 'თა_' | 'საა' | 'საე' | 'საი' | 'საო' | 'საუ' | 'სამ' | 'თან' | 'სახ' |
| 'საკ' | 'სე_' | 'შეა' | 'შეე' | 'შეი' | 'თეო' | 'თეუ' | 'შემ' | 'თელ' | 'შეს' |
| 'შეკ' | 'ში_' | 'ცია' | 'ციე' | 'ციი' | 'სიო' | 'სიუ' | 'სიმ' | 'სინ' | 'თის' |
| 'სიკ' | 'თო_' | 'სოა' | 'თოე' | 'თოი' | 'ციო' | 'შოუ' | 'თობ' | 'შორ' | 'თოქ' |
| 'ციოტ' | 'თუ_' | 'შუა' | 'სუე' | 'თუი' | 'თუო' | 'თუუ' | 'თუმ' | 'თულ' | 'სუხ' |
| 'თუკ' | 'სზ_' | 'თზა' | 'ჩზე' | 'თზი' | 'თზო' | 'ცზუ' | 'სზრ' | 'სგ_' | 'სგა' |
| 'სგე' | 'სგი' | 'თგუ' | 'თგვ' | 'თგზ' | 'სღ_' | 'ცღა' | 'ცღე' | 'ცღი' | 'ცღო' |
| 'ცღუ' | 'სღვ' | 'სღრ' | 'თვ_' | 'თვა' | 'თვე' | 'თვი' | 'შვო' | 'შვნ' | 'თვს' |
| 'ჩვწ' | 'თზა' | 'თზე' | 'სზო' | 'სთ_' | 'სთა' | 'სთე' | 'სთი' | 'სთუ' | 'სთვ' |
| 'სთნ' | 'სთხ' | 'სკ_' | 'შკა' | 'სკე' | 'სკი' | 'სკო' | 'სკუ' | 'სკვ' | 'სკნ' |

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|--------|-------|-------|-------|
| 'სკო' | 'თლ_' | 'თლა' | 'თლე' | 'თლი' | 'თლო' | 'თლუ' | 'თლმ' | 'შლდ' | 'თლშ' |
| 'თლწ' | 'სმ_' | 'თმა' | 'თმე' | 'სმი' | 'სმო' | 'სმუ' | 'სმგ' | 'ცმლ' | 'თმც' |
| 'თმყ' | 'ცნ_' | 'შნა' | 'ჩნე' | 'ცნი' | 'ცნო' | 'შნუ' | 'შნვ' | 'ჩნდ' | 'შნთ' |
| 'სპ_' | 'სპა' | 'სპე' | 'სპი' | 'სპო' | 'სპუ' | 'სპრ' | 'თრ_' | 'თრა' | 'თრე' |
| 'თრი' | 'შრო' | 'სრუ' | 'ჩრდ' | 'თრთ' | 'ცრტ' | 'თს_' | 'თსა' | 'თსი' | 'სსუ' |
| 'სსვ' | 'სსს' | 'სტ_' | 'სტა' | 'სტე' | 'სტი' | 'სტო' | 'სტუ' | 'სტვ' | 'სტრ' |
| 'სტშ' | 'სტყ' | 'სფა' | 'სფე' | 'თფი' | 'სფო' | 'სფუ' | 'სფრ' | 'შქ_' | 'ჩქა' |
| 'სქე' | 'თქი' | 'თქო' | 'შქუ' | 'თქვ' | 'თქლ' | 'შქთ' | 'სლა' | 'სლი' | 'სყა' |
| 'სყი' | 'სყო' | 'თყვ' | 'თშ_' | 'თშა' | 'თშე' | 'თში' | 'თშთ' | 'სშტ' | 'სჩ_' |
| 'შჩა' | 'სჩე' | 'სჩვ' | 'თჩრ' | 'სჩქ' | 'თც_' | 'სცა' | 'სცე' | 'სცი' | 'სცო' |
| 'სცვ' | 'სცდ' | 'სცქ' | 'სძა' | 'სძე' | 'სძი' | 'სძვ' | 'სძლ' | 'სწა' | 'სწე' |
| 'სწი' | 'სწო' | 'სწუ' | 'სწვ' | 'სწლ' | 'სწყ' | 'სჭა' | 'სჭი' | 'სჭვ' | 'სხ_' |
| 'ცხა' | 'თხე' | 'თხი' | 'ცხო' | 'თხუ' | 'სხვ' | 'ცხლ' | 'თხს' | 'თხტ' | 'სჯა' |
| 'სჯე' | 'სჯი' | 'სჯო' | 'სჯუ' | 'სჯვ' | 'სჯდ' | 'ცჰა' | 'ცჰე' | 'შჰო' | 'შჰუ' |
| 'ს_ა' | 'ს_ე' | 'ს_ი' | 'ს_ო' | 'ს_უ' | 'ს_მ' | 'ს_დ' | 'ს_ს' | 'ს_კ' | 'კა_' |
| 'კაა' | 'ყაე' | 'კაი' | 'ყაო' | 'ყაუ' | 'კავ' | 'კარ' | 'კას' | 'კაპ' | 'კე_' |
| 'კეა' | 'ყეე' | 'პეი' | 'ყეო' | 'კეუ' | 'კებ' | 'პერ' | 'კეთ' | 'პეტ' | 'კი_' |
| 'კია' | 'კიე' | 'კიი' | 'კიო' | 'კიუ' | 'კიბ' | 'პირ' | 'კით' | 'პიტ' | 'ყო_' |
| 'კოა' | 'პოე' | 'კოი' | 'კოო' | 'პოუ' | 'ყოვ' | 'კონ' | 'ყოფ' | 'კოტ' | 'კუ_' |
| 'პუა' | 'ყუე' | 'ყუი' | 'კუმ' | 'კუნ' | 'კუთ' | 'კუტ' | 'ყბა' | 'ყბე' | 'კბი' |
| 'კბო' | 'კბლ' | 'კბი' | 'ყდა' | 'ყდე' | 'ყდი' | 'კდო' | 'კდუ' | 'კდბ' | 'კდრ' |
| 'კვ_' | 'ყვა' | 'ყვე' | 'კვი' | 'ყვო' | 'კვლ' | 'კვთ' | 'ყვტ' | 'პზე' | 'კთა' |
| 'პთმ' | 'კთხ' | 'კკლ' | 'კლა' | 'კლე' | 'ყლი' | 'კლო' | 'კლუ' | 'კლვ' | 'კლდ' |
| 'პლს' | 'კმა' | 'კმე' | 'ყმი' | 'ყმუ' | 'კმლ' | 'ყნა' | 'ყნე' | 'ყნი' | 'ყნო' |
| 'კნუ' | 'ყნდ' | 'კნჭ' | 'კრა' | 'კრე' | 'კრი' | 'პრო' | 'კრუ' | 'ყრმ' | 'ყრდ' |
| 'კრთ' | 'კს_' | 'კსა' | 'კსე' | 'კსი' | 'პსო' | 'პსუ' | 'კსვ' | 'კსნ' | 'კსს' |
| 'კტ_' | 'პტა' | 'პტე' | 'კტი' | 'პტო' | 'პტუ' | 'პტმ' | 'პტნ' | 'პყა' | 'პყრ' |
| 'პშე' | 'კში' | 'ყწი' | 'კხე' | 'კხი' | 'კ_ა' | 'კ_ი' | 'პ_ო' | 'კ_უ' | 'ყ_ვ' |
| 'პ_რ' | 'კ_შ' | 'პ_ტ' | 'ხა_' | 'ქაა' | 'ჰაე' | 'ქაი' | 'ხაო' | 'ქაუ' | 'ხავ' |
| 'ქარ' | 'ფას' | 'ხატ' | 'ხე_' | 'ხეა' | 'ხეე' | 'ხეი' | 'ხეო' | 'ხეუ' | 'ხებ' |
| 'ხელ' | 'ფეს' | 'ფეტ' | 'ხი_' | 'ფია' | 'ფიე' | 'ფიი' | 'ფიო' | 'ფიუ' | 'ხივ' |
| 'ფილ' | 'ფიქ' | 'ფიკ' | 'ფო_' | 'ხოა' | 'ხოე' | 'ფოი' | 'ხოო' | 'ხოუ' | 'ხოვ' |
| 'ხოლ' | 'ქოს' | 'ფოტ' | 'ქუ_' | 'ხუა' | 'ქუე' | 'ხუი' | 'ხუმ' | 'ხურ' | 'ხუთ' |
| 'ფუტ' | 'ჰბე' | 'ხბი' | 'ხბო' | 'ჰგა' | 'ჰგი' | 'ჰგო' | 'ჰგვ' | 'ჰგლ' | 'ფდ_' |
| 'ხდა' | 'ხდე' | 'ხდი' | 'ხდო' | 'ხდუ' | 'ხდნ' | 'ხდს' | 'ხვ_' | 'ხვა' | 'ქვე' |
| 'ხვი' | 'ხვო' | 'ხვმ' | 'ხვდ' | 'ქვს' | 'ხზე' | 'ხზეუ' | 'ხზვ' | 'ხთ_' | 'ხთა' |
| 'ფთი' | 'ქთო' | 'ხთუ' | 'ფთვ' | 'ფთრ' | 'ჰკა' | 'ჰკი' | 'ხკუ' | 'ჰკვ' | 'ჰკრ' |
| 'ფვს' | 'ხლ_' | 'ხლა' | 'ხლე' | 'ხლი' | 'ხლო' | 'ხლუ' | 'ფლმ' | 'ხლდ' | 'ხლს' |
| 'ხლტ' | 'ქმ_' | 'ხმა' | 'ქმე' | 'ქმი' | 'ხმო' | 'ქმუ' | 'ქმნ' | 'ქმფ' | 'ქნა' |
| 'ქნე' | 'ხნი' | 'ქნო' | 'ქნბ' | 'ქნზ' | 'ხნს' | 'ჰჰა' | 'ჰჰე' | 'ჰჰრ' | 'ფრა' |
| 'ხრე' | 'ხრი' | 'ფრო' | 'ფრუ' | 'ქრბ' | 'ქრდ' | 'ფროთ' | 'ხრტ' | 'ხს_' | 'ქსა' |
| 'ხსე' | 'ქსი' | 'ქსო' | 'ქსუ' | 'ხსნ' | 'ქსთ' | 'ფსკ' | 'ქტ_' | 'ქტა' | 'ქტე' |
| 'ქტი' | 'ქტო' | 'ქტუ' | 'ქტმ' | 'ქტრ' | 'ქტს' | 'ჰფე' | 'ჰფი' | 'ფქა' | 'ჰქო' |
| 'ჰქვ' | 'ჰქრ' | 'ჰქც' | 'ჰღვ' | 'ჰყა' | 'ჰყე' | 'ჰყო' | 'ჰყუ' | 'ჰყვ' | 'ჰყრ' |

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 'ფმ_' | 'ფმა' | 'ქმე' | 'ხმი' | 'ხმო' | 'ხმუ' | 'ფმვ' | 'ფმს' | 'ფმტ' | 'ქჩე' |
| 'ხჩი' | 'ქცა' | 'ქცე' | 'ქცი' | 'ქცო' | 'ქცნ' | 'ხჭა' | 'ჰჭე' | 'ფხ_' | 'ფხა' |
| 'ფხე' | 'ფხი' | 'ფხო' | 'ფხუ' | 'ფხვ' | 'ფხდ' | 'ხჯე' | 'ფჯგ' | 'ხ_ა' | 'ქ_ე' |
| 'ქ_ი' | 'ქ_ო' | 'ქ_უ' | 'ხ_მ' | 'ქ_დ' | 'ხ_კ' | 'ხ_პ' | 'ტა_' | 'წაა' | 'წაე' |
| 'ტაი' | 'ტაო' | 'წაუ' | 'წამ' | 'წარ' | 'ტაც' | 'ტატ' | 'ტე_' | 'წეა' | 'წეე' |
| 'ტეი' | 'წეო' | 'წეუ' | 'ტებ' | 'ტერ' | 'ტეს' | 'ტეტ' | 'ტი_' | 'ტია' | 'წიე' |
| 'ტიი' | 'ტიო' | 'ტიუ' | 'წიგ' | 'წინ' | 'ტის' | 'ტიკ' | 'ტო_' | 'ტოა' | 'ტოე' |
| 'ტოი' | 'ტოო' | 'ჭოუ' | 'ტომ' | 'ტორ' | 'ტოქ' | 'ტოპ' | 'ტუ_' | 'ტუა' | 'ტუე' |
| 'ტუი' | 'ტუმ' | 'ტურ' | 'წუთ' | 'ტუტ' | 'ტბა' | 'ტბე' | 'ტბი' | 'ტბუ' | 'ტდ_' |
| 'ჭდა' | 'ჭდე' | 'ჭდი' | 'ჭდო' | 'ჭდვ' | 'ტდნ' | 'წვ_' | 'წვა' | 'წვე' | 'ტვი' |
| 'ჭვო' | 'ჭვუ' | 'ჭვგ' | 'წვდ' | 'ჭვს' | 'წვყ' | 'ტზე' | 'ტთ_' | 'ტთა' | 'ტკა' |
| 'ტკე' | 'ტკი' | 'ტკო' | 'ტკუ' | 'ტკვ' | 'წკრ' | 'ტლ_' | 'ტლა' | 'წლე' | 'წლი' |
| 'წლო' | 'ტლუ' | 'ტლდ' | 'ტლნ' | 'წლს' | 'ტმა' | 'წმე' | 'წმი' | 'წმო' | 'წმუ' |
| 'წმვ' | 'წმდ' | 'ტმს' | 'წმყ' | 'წნ_' | 'ტნა' | 'წნე' | 'ტნი' | 'ტნო' | 'ტნუ' |
| 'ტნმ' | 'ტნს' | 'ტპო' | 'ტრ_' | 'ტრა' | 'ტრე' | 'ტრი' | 'ტრო' | 'ტრუ' | 'ტრმ' |
| 'ტრზ' | 'ტრს' | 'ტს_' | 'ტსა' | 'ტსო' | 'ტტე' | 'ტფი' | 'ტქა' | 'ტქი' | 'ტყ_' |
| 'წყა' | 'წყე' | 'წყი' | 'წყო' | 'ტყუ' | 'ტყვ' | 'წყლ' | 'ტში' | 'ტჩი' | 'ტცა' |
| 'ტწი' | 'ტხა' | 'ტხე' | 'ტჰა' | 'ტჰე' | 'ტ_ა' | 'ტ_ე' | 'ტ_ი' | 'ტ_ო' | 'ტ_უ' |
| 'ტ_გ' | 'ტ_ნ' | 'ტ_ს' | 'ტ_კ' | | | | | | |

წარმომადგენელთა ბაზა RepCL შედგება მესამე ცხრილში მოყვანილი ტრიფონებისაგან.

შემდეგი ეტაპია ტრიფონების ცხრილზე დაყრდნობით ჩასაწერი ტექსტის შედგენა.

6. აკუსტიკური ბაზის ჩაწერა.

აკუსტიკური ბაზის ერთეულთა მისაღებად საჭიროა შევადგინოთ და ჩავწეროთ წარმომადგენელთა ბაზის შემცველი ტექსტი.

ჩასაწერი ტექსტი უმჯობესია შედგებოდეს დასრულებული წინადადებებისაგან. იგი უნდა შეიცავდეს წარმომადგენელთა ბაზის ყველა ელემენტს.

ამ ეტაპზე ჩვენ მივიღეთ გადაწყვეტილება, ტექსტი შეგვედგინა ცალკეული, წარმომადგენელთა ბაზის შემცველი სიტყვებისაგან. ამ გადაწყვეტილების მიზეზი ძირითადად ორია: პირველი - ჩვენ არ განვიხილავთ პუნქტუაციას, ვინაიდან სასვენი ნიშნების გავლენა მეტყველებაზე მოითხოვს დამატებით კვლევებს. მეორე - წინადადებები შეიცავენ დიდი რაოდენობით "არასაჭირო" ალოფონებს. ჩანაწერიდან საჭირო ერთეულების ამოღება მოითხოვს დიდ დროს. ვიმედოვნებთ, შემდგომში გვექნება საშუალება, უფრო სრულყოფილი ტექსტის ჩაწერისა, რომელშიც გათვალისწინებული იქნება წინამდებარე კვლევების შედეგებიც.

ტექსტი შევადგინეთ შემდეგი წესით: წარმომადგენელთა ბაზის პირველი ელემენტისათვის ტექსტიდან ამოვიღეთ მისი შემცველი სიტყვა. მიღებულ სიტყვაში მოვძებნეთ წარმომადგენელთა ბაზის სხვა ელემენტებიც, რომლებსაც შეიცავდა აღნიშნული სიტყვა. წარმომადგენელთა ბაზას გამოვაკელით "მონიშნული" ტრიფონები.

დარჩენილი ბაზის პირველი ელემენტის მიმართ განვახორციელეთ ანალოგიური ოპერაციები და მივიღეთ მეორე სიტყვა და შემცირებული ბაზა... და ასე შემდეგ, ბაზის ამოწურვამდე. შემდეგ სიტყვათა სიმრავლიდან ამოვავდეთ გამეორებული სიტყვები, თუ ასეთები გვქონდა. მიღებული სიტყვათა სიმრავლე დავალაგეთ სიგრძის (სიმბოლოებში) მატების რიგით. მიღებული, ცალკეული სიტყვებისაგან შედგენილი ტექსტი დავბეჭდეთ.

აღწერილი ალგორითმი ჩასაწერი ტექსტის შედგენისა, რეალიზებულია m პროგრამულ ენაზე. პროგრამის გამოყენებით მივიღეთ 2360 სიტყვის შემცველი ტექსტი. უნდა აღვნიშნოთ, რომ ტექსტი შეიცავს მცირეოდენ შეცდომებს (სიტყვას აკლია ან აქვს ზედმეტი რაიმე სიმბოლო, ან, შემთხვევით, სხვა სიმბოლოა დაბეჭდილი და სხვა), მაგრამ ეს არ ქმნის რაიმე პრობლემას, ვინაიდან ჩვენ გვაინტერესებს არა თავად სიტყვები, არამედ ასოთა შეთანხმებები.

ჩაწერა განვახორციელეთ საქართველოს საპატრიარქოსთან არსებულ კოლხური კულტურის ცენტრის ხმის ჩამწერ სტუდიაში კომპოზიტორ მერაბ მამულაშვილის დახმარებით. ტექსტი წაიკითხა ქ-ნ ფატი თურმანიძემ. ჩანაწერის ფორმატია 44100 კილოჰერცი, 16 ბიტი, მონო. ბ-ნმა მერაბ მამულაშვილმა აუდიო ფაილი დაჭრა ცალკეულ სიტყვებად, რის შემდეგაც გადმოგვეცა ჩვენ კომპაქტდისკზე ჩაწერილი ტრეკების (TR-00001.wav, TR-00002.wav, და ა. შ) სახით. ყოველი ტრეკი შეიცავს ერთ სიტყვას.

ტექსტის და ჩანაწერის სინქრონიზაციის და მათში ორიენტაციის გაადვილების მიზნით ტრეკებს შევუცვალეთ სახელები და შესაბამის ფაილებს დავარქვით ის სიტყვა, რომელიც მასში წერია. მაგ. სიტყვას "გამარჯობა" შეესაბამება "gamarjoba.wav". მათი ურთიერთგადაყვანისათვის გვაქვს ორი მარტივი m-ფუნქცია **word = fn2word(filename)** და **filename = word2fn(word)**.

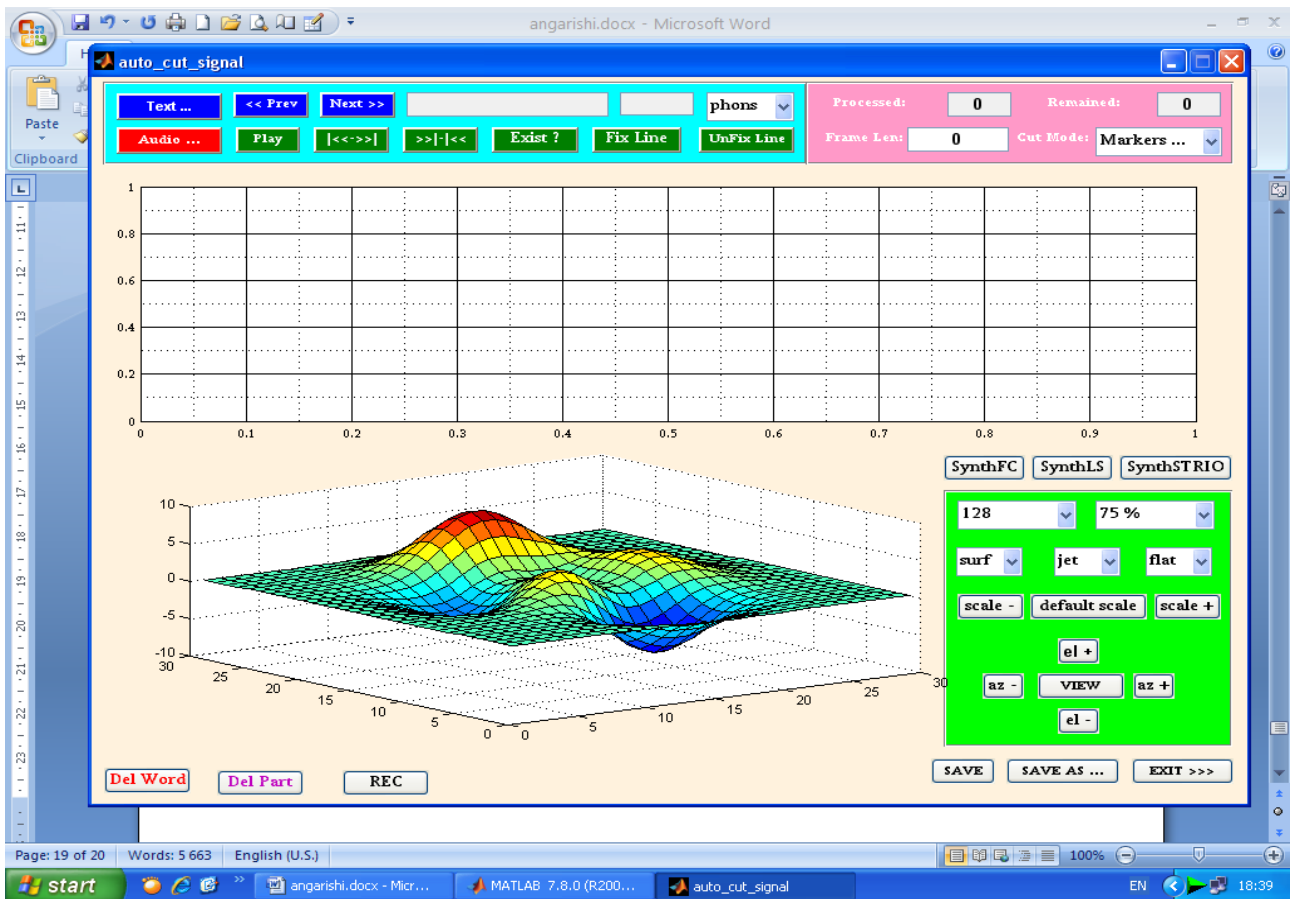
7. სიტყვების სეგმენტაცია აკუსტიკურ ერთეულებად.

ერთ-ერთი მნიშვნელოვანი საკითხია სიტყვების სეგმენტაცია აკუსტიკურ ერთეულებად. ჩვენს შემთხვევაში საჭიროა ზუსტი საზღვრების გავლება ალოფონთა შორის. სეგმენტაციის სიზუსტეზე არის დამოკიდებული სინთეზირებული ხმის სისუფთავე ზედმეტი ჩქამებისაგან.

აღნიშნული მიზნით, MatLab-ის გარემოში GUIDE გამოყენებით შევქმენით გრაფიკული ინტერფეისის მქონე პროგრამა **auto_cut_signal**. მისი საწყისი გრაფიკული ფანჯარა მოცემულია მეოთხე სურათზე.

აღნიშნულ პროგრამას აქვს ბევრი სასარგებლო ფუნქცია. აღწერთ მოლოდ მათ ერთ ნაწილს. ყოველ ღილაკზე მაუსის შეჩერებით ვღებულობთ ღილაკის დანიშნულებას დროებით ფანჯარაში, ამიტომ ღილაკების აღწერას არ გთავაზობთ.

ვინაიდან პროგრამა შედგენილია უშუალოდ ჩვენს წინაშე მდგარი ამოცანების გადასაჭრელად, იგი მაქსიმალურად არის მორგებული მათზე.



სურ. 4. პროგრამა **auto_cut_signal** საწყისი ფანჯარა.

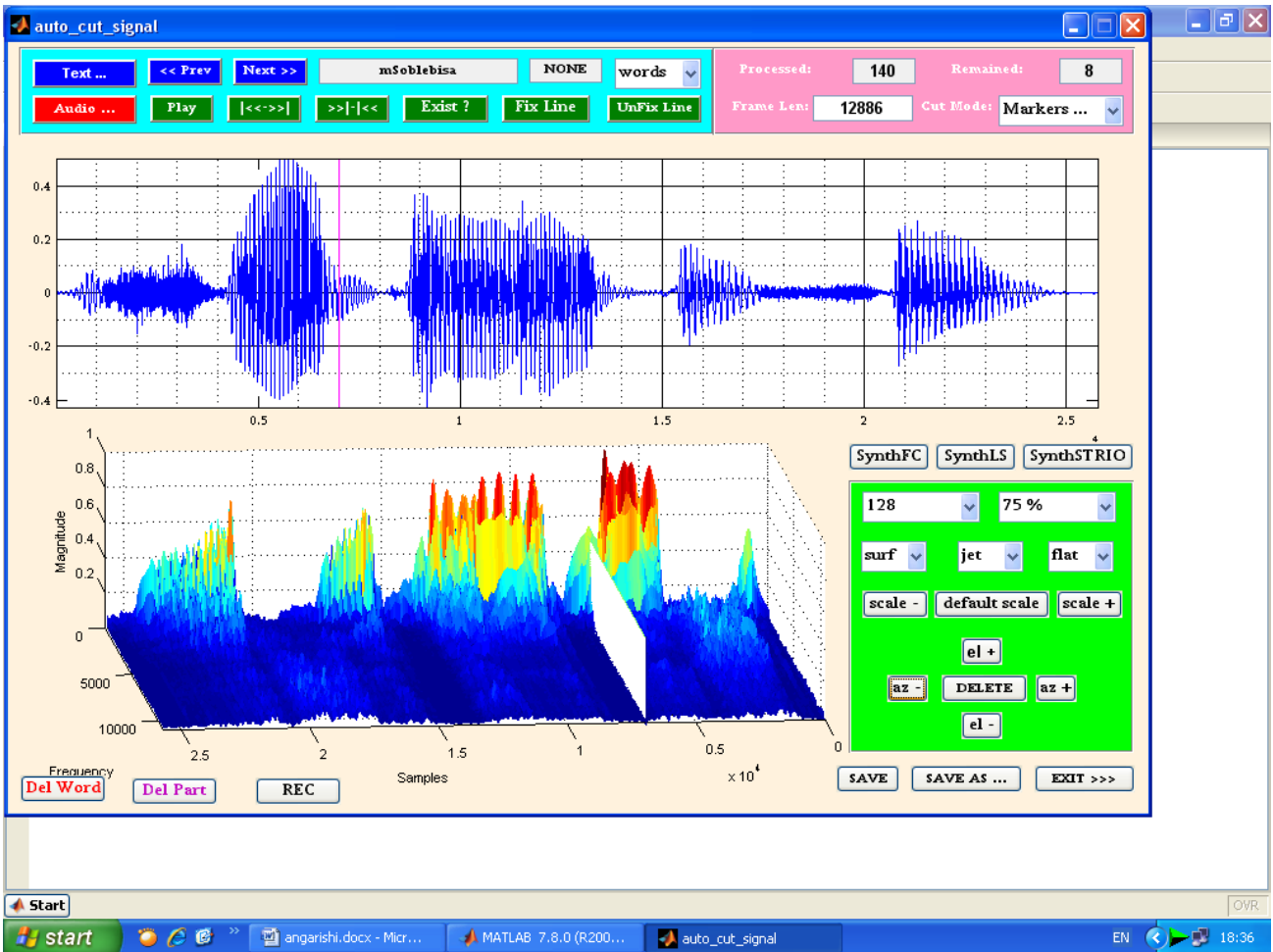
ვთქვათ გვინდა ჩაწერილი ტექსტში შემავალი სიყცეების სეგმენტაცია. ამისათვის ნაწილებს ასარჩევ ჩამოსაშლელ მენიუში ვირჩევთ პუნქტს - “**words**”, ხოლო დაჭრის ტიპის ასარჩევ მენიუში - “**Markers Only**”. შემდეგ მაუსით ვაწკაპუნებთ ღილაკზე **Text...** გაიხსნება ფაილის არჩევის სტანდარტული ფანჯარა, სადაც ავირჩევთ ჩვენი ტექსტის შემცველ ფაილს.

პროგრამა წაიკითხავს მას და შექმნის სიტყვების მასივს. გრაფიკული ეკრანის შესაბამის ველში (პირველი სტრიქონის მესამე ღილაკის შემდეგ) გამოიტანს პირველ სიტყვას. ავტომატურად მოძებნის ამ სიტყვის შესაბამის აუდიო ფაილს და გამოიტანს ზემოთა გრაფიკულ ფანჯარაში. ღილაკზე **Play** დაწკაპუნებით შეგვიძლია მოვისმინოთ ჩანაწერი.

სიგნალის ფანჯარაში საჭირო ადგილას მაუსის დაწკაპუნებით მასზე ჩნდება იასამნისფერი მარკერი (სურ. 5). თუ ამ დროს დავაწკაპუნებთ ღილაკზე **Play**, მოვისმენთ სიგნალის მონაკვეთს დასაწყისიდან იასამნისფერ მარკერამდე. მცირე პაუზის (2 წამი) შემდეგ ავტომატურად მოვისმენთ მონაკვეთს მარკერიდან ბოლომდე. დამაკმაყოფილებელი შედეგის შემთხვევაში დავაწკაპუნებთ ღილაკზე **Fix Line**, რის შედეგადაც მარკერი დაფიქსირდება და მიიღებს წითელ ფერს.

როდესაც მოსმენით ძნელია ზუსტი საზღვრის დადგენა, მარკერის დაფიქსირებამდე დავაწკაპუნოთ მწვანე პანელზე განთავსებულ ღილაკზე **VIEW**. ქვედა

გრაფიკულ ფანჯარაში გამოჩნდება სიგნალის სამგანზომილებიანი სპექტროგრამა (წარწერა **VIEW** შეიცვლება წარწერით **DELETE**). სპექტროგრამა გაყოფილი იქნება თეთრი სიბრტყით ჩვენი მარკერის მდებარეობის შესაბამისად. საშუალება გვებძლება გამოსახულება ვატრიალოთ სივრცეში და ვიზუალურად დავაკვირდეთ მარკერის მდებარეობას დროით-სპექტრულ სივრცით გამოსახულებაზე. დაზუსტების შემდეგ დავაფიქსირებთ მარკერს.



სურ. 5. პროგრამა **auto_cut_signal** მოქმედებაში.

ანალოგიური გზით დავსვამთ და დავაფიქსირებთ ყველა საჭირო მარკერს სიგნალზე. შემდეგ ქვემოთ განთავსებულ დილაკზე **SAVE** დაწკაპუნებით პროგრამა ავტომატურად გადაიყვანს დასმული მარკერების კოორდნატებს სამპლებში და შეინახავს სახელით **sityva.csv**. **sityva** აქ შეიცვლება იმ სიტყვის ლათინური ტრანსკრიფციით, რომელიც მუშავდება.

პროგრამა თვალყურს ადევნებს დასამუშავებელ სიტყვაში სიმბოლოების რაოდენობას და დასმული მარკერების რაოდენობას. მხოლოდ მათი შესაბამისობის შემთხვევა შეასრულებს ჩაწერის ოპერაციას და გრაფიკული ფანჯრიდან გააქრობს ყველა მარკერს. წინააღმდეგ შემთხვევაში ჩაწერა არ მოხდება, მარკერები ადგილზე დარჩებიან და მოგვეცემა შესწორების საშუალება. ამით ბაზა დაცულია ყალბი მონაცემებისაგან.

შემდეგ სიტყვაზე გადასასვლელად ვაწკაპუნებთ ღილაკზე **Next**. წინა სიტყვაზე დასაბრუნებლად ვიყენებთ ღილაკს **Prev**.

პროგრამიდან გამოსასვლელად **უნდა გამოვიყენოთ** ღილაკი **Exit**. ასეთ შემთხვევაში პროგრამა ხურავს აქტიურ ობიექტებს და იმახსოვრებს დამუშავებული სიტყვების რაოდენობას. შემდეგი სეანსი გაგრძელდება დამახსოვრებული სიტყვიდან. სხვა გზით გამოსვლის შემთხვევაში ეს არ მოხდება.

შესაძლებელია ვცვალოთ სიგნალის გამოსახულების მასშტაბი, სპექტროგრამის პარამეტრები, ვატრიალოთ სივრცეში სიგნალის სამგანზომილებიანი გამოსახულება, ვცვალოთ ფერები და სხვა.

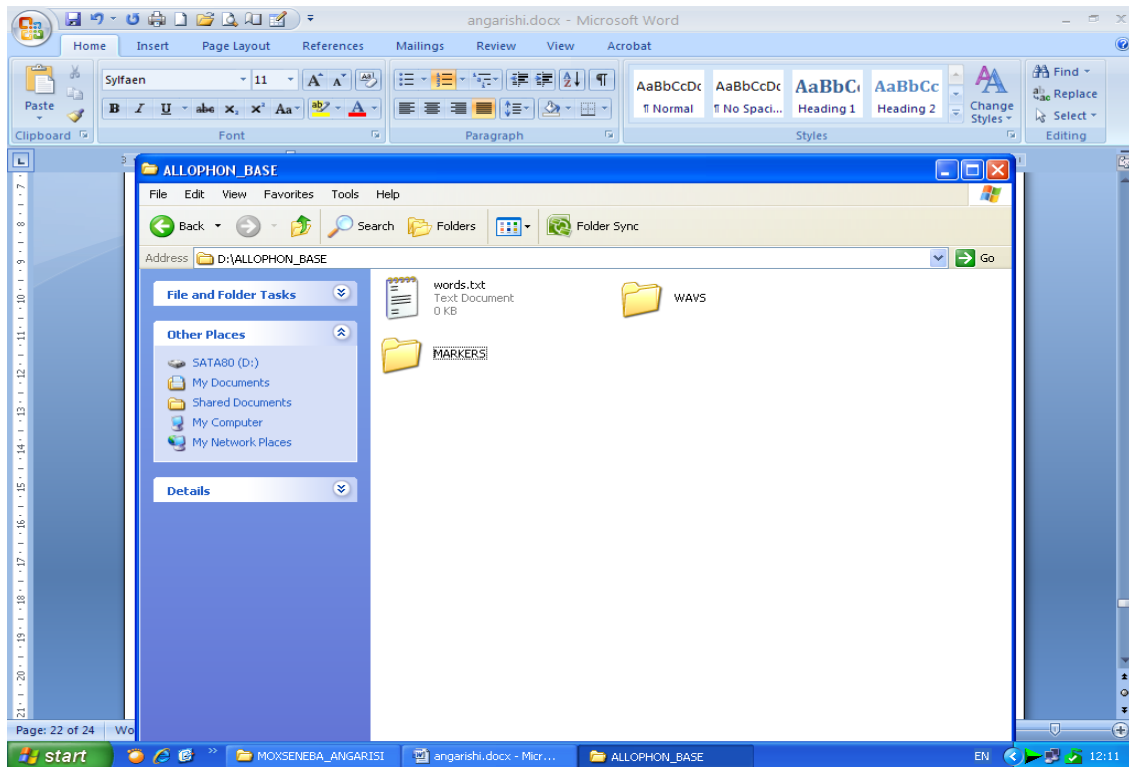
რაც მთავარია, ჩაწერის შემდეგ **SynthX** დაწკაპუნებით გვაქვს შესაძლებლობა მოვისმინოთ სინთეზი დასამუშავებელი სიტყვისა. ეს გვამლევს საშუალებას დროულად, მარკირების პროცესში გავასწოროთ შეცდომები. პროგრამაში შესაძლებელია სხვადასხვა სახის სინთეზატორის მოდელების ინტეგრაცია, რომელთა ტესტირებასაც ვაპირებთ. ამიტომ შეგვიძლია ერთდროულად მოვუსმინოთ ყველა მათგანს და შევადაროთ.

არის სხვა შესაძლებლობებიც. პროგრამა არის ღია ტექსტით და საჭიროების შემთხვევაში შესაძლებელია მასში საჭირო ცვლილებების და დამატებების მარტივად შეტანა.

8. სინთეზი და ტესტირება.

ყოველი სიტყვის მარკირების შემდეგ ჩვენ გვექნება ბაზის სამი შემადგენელი კომპონენტი:

1. ჩაწერილი სიტყვებისაგან შედგენილი ტექსტი **words.txt**;
2. სადალდე **WAVS**, რომელშიც ინახება აუდიო ფაილები - **word.wav** სიტყვების შესაბამისი სახელებით;
3. საქარალდე **MARKERS**, სადაც ინახება შესაბამისი სახელით **word.csv** სიტყვის მარკერების შემცველი ფაილები.



სურ. 6. საქარალდე ALLOPHONE_BASE სამი კომპონენტით.

ეს სამი კომპონენტი საკმარისია ვიცოდეთ რა ალოფონები გვაქვს ბაზაში და საჭიროების შემთხვევაში ამოვიღოთ ყოველი მათგანი.

ქვემოთ მოყვანილია ბაზის ტესტირების m-პროგრამის ერთი ფრაგმენტი, სადაც ილუსტრირებულია შესაბამისი ალგორითმი. მწვანე ფერით გამოყოფილია კომენტარები.

```
function signal = playtext(txt)
```

```
%=====
%პროგრამა ალოფონურიბაზისტესტირებისათვის
%=====
```

```
% თუ არგუმენტი ცარიელია, გავხსნათ ფაილის არჩევის სტანდარტული ფანჯარა
```

```
if nargin < 1
    [fnT, fpT] = uigetfile('*.txt', 'Get File Name and Path');
if (length([fnT fpT]) == 2)
    fnT = [];
    fpT = [];
end
```

```
if isempty([fpT fnT])
return;
end
```

```
txt = mytextscanu([fpT fnT]);
end
```

```
% მონაცემთა აღწერა და ინიციალიზაცია
```

```

global nL nR nL1 nR1 nL2 nR2 nL3 nR3 nL4 nR4
global LB RB LB1 RB1 LB2 RB2 LB3 RB3 LB4 RB4
global TRIO SIMTRIO1 SIMTRIO2 SIMTRIO3 SIMTRIO4
global triodata simtriodata1 simtriodata2 simtriodata3 simtriodata4
global simtest
    simtest = ' ';

    load 'projtest'nLnRnL1nR1nL2nR2nL3nR3nL4nR4;
    load 'projtest'LBRBLB1RB1LB2RB2LB3RB3LB4RB4;
    load 'projtest'TRIOSIMTRIO1SIMTRIO2SIMTRIO3SIMTRIO4;
    load 'projtest'triodatasimtriodata1simtriodata2simtriodata3simtriodata4;
    fs = 22050;
    comma = zeros(1, 3000);
    pnt = zeros(1, 6000);

    % ტექსტის ნორმალიზაცია.
    txt = str_normalization(txt);

    % ტექსტის დაყოფა წინადადებებად.
    [sent, nsent] = get_sentens(txt);

    % ციკლი წინადადებების მიხედვით.
    for ii = 1:nsent
        % მიმდინარე წინადადების დაყოფა მარტივ წინადადებებად.
        [subsent, nsubsent] = get_subsentens(sent{ii});

        % ციკლი მარტივი წინადადებების მიმართ.
        for jj = 1:nsubsent
            % გავასუფთავოთ მარტივი წინადადება არაკითხვადი სიმბოლოებისაგან.
            subs = strtonewstr3(subsent{jj});
            % დავყოთ მარტივი წინადადება სიტყვებად.
            [wrд, nwrд] = strtowords(subs);
            % ციკლი ცალკეული სიტყვების მიმართ.
            for kk = 1:nwrд
                %დავყოთ სიტყვა ტრიფონებად.
                [trio, ntrio] = get_trio(wrd{kk});
                if (ntrio == 0)
                    return;
                end

                signal = zeros(1, 100);

            % ციკლი ყოველი ტრიფონის მიმართ.
            for mm = 1:ntrio

```

```

        x = trio{mm};
if x(2) ~= '_'
% ამოვიღოთ ბაზიდან მიმდინარე ალოფონის მონაცემები.
[ad, nad] = get_allophondata(x);
% დავსვათ სიგნალის საშუალო ნულზე
ad = zero_aver(ad);
if (nad == 0)
% თუ არ გვაქვს ალოფონი გამოვიტანოთ შეტყობინება და შევწყვიტოთ
% ტესტირება.
        display(strcat('არ მაქვს ალოფონი', '---', x));
return;
end

% განვსაზღვროთ მომდინარე სიმბოლო ხმოვანია თუ თანხმოვანი.
        t = get_type(trio{mm}(2));

% აუდიო სტრიქონის მიმდინარე სეგმენტის ფორმირება.
if (t == 1)% ხმოვანია
if simtest(end) == '0'% მიმდინარე ხმოვანი ეკუთვნის ზუსტთა ბაზას.
ad = makesoundlen(ad, 2400, 'B', 'L', 45);
else% მიმდინარე ხმოვანი არ ეკუთვნის ზუსტთა ბაზას.
ad = makesoundlen(ad, 2400, 'B', 'L', 1);
end
elseif (t > 1)% თანხმოვანია
if simtest(end) == '0'% მიმდინარე თანხმოვანი ეკუთვნის ზუსტთა ბაზას.
ad = makesoundlen(ad, 2000, 'E', 'R', 45);
else% მიმდინარე თანხმოვანი არ ეკუთვნის ზუსტთა ბაზას.
ad = makesoundlen(ad, 2000, 'E', 'R', 5);
end
end
end
% მიმდინარე სეგმენტის მიერთება ძირითად სტრიქონთან.
        signal = signalplus(signal, ad);
else
% სიტყვათაშორისი ინტერვალის ფორმირება.
        signal = [signal zeros(1, 200)];
end
end

% სიტყვის ხმის სიმაღლის განსაზღვრა.
        signal = normalize_snd(signal, 0.4);
% სიტყვის პროფილის ფორმირება.
        signal = formgenerator2(signal, 41, 0.5);
% სიტყვის წარმოთქმა.
        wavplay(signal, fs);
end

```

% პაუზა მძიმესთან.

wavplay(comma, fs);

end

% პაუზა წერტილთან.

wavplay(pnt, fs);

end

*% პროგრამის ეს მონაკვეთი ითვლის რამდენი ალოფონი გვექონდაზუსტ ალოფონთა
ბაზაში და % რამდენი სხვა ბაზებიდან. ეკრანზე გამოაქვს საერთო შედეგი
პროცენტებში.*

lsim = length(simtest);

nt = findstr('0', simtest);

if isempty(nt)

BAZE0 = 0;

else

BAZE0 = length(nt);

end

*str0 = ['BAZE0 = ' int2str(round(100*BAZE0/lsim)) '%'];*

disp(str0);

nt = findstr('1', simtest);

if isempty(nt)

BAZE1 = 0;

else

BAZE1 = length(nt);

end

*str0 = ['BAZE1 = ' int2str(round(100*BAZE1/lsim)) '%'];*

disp(str0);

nt = findstr('2', simtest);

if isempty(nt)

BAZE2 = 0;

else

BAZE2 = length(nt);

end

*% str0 = ['BAZE2 = ' int2str(round(100*BAZE2/lsim)) '%'];*

% disp(str0);

nt = findstr('3', simtest);

if isempty(nt)

BAZE3 = 0;

else

BAZE3 = length(nt);

```

end
% str0 = ['BASE3 = ' int2str(round(100*BAZE3/lsim)) '%'];
% disp(str0);

nt = findstr('4', simtest);
if isempty(nt)
    BAZE4 = 0;
else
    BAZE4 = length(nt);
end
% str0 = ['BASE4 = ' int2str(round(100*BAZE4/lsim)) '%'];
% disp(str0);

str0 = ['BASE2 = ' int2str(round(100*(BAZE2+BAZE3+BAZE4)/lsim)) '%'];
disp(str0);

simtest(1:end) = [];

end

```

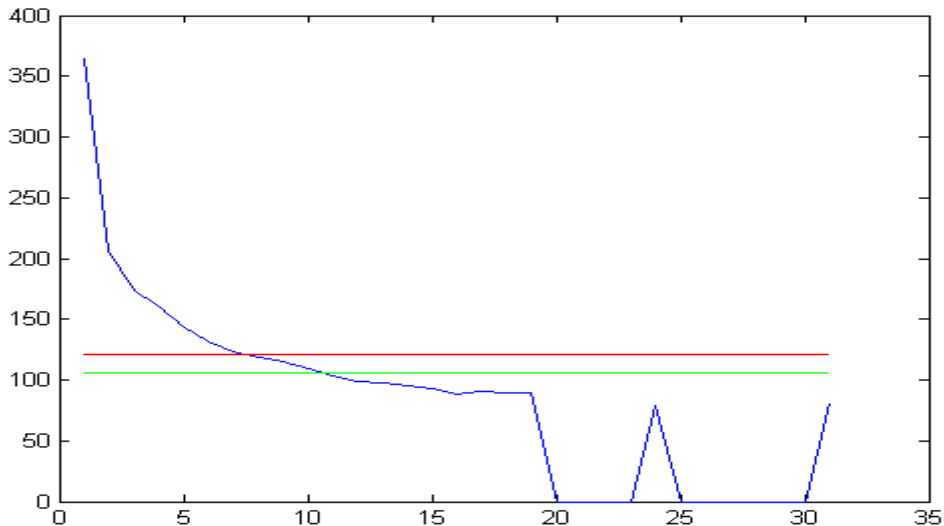
ტესტირების პროგრამას აქვს მარტივი ალგორითმი. მისი ფუნქციაა არა მოგვცეს სრულფასოვანი სუნთეზი, არამედ დაგვცემარის სრულფასოვანი ბაზის შედგენაში. მისი გამოყენებით ვისმენტ სინთეზს, ისეთს, "როგორც არის". ეს საშუალებას გვაძლევს შევნიშნოთ მარკირების უზუსტობა.

პროგრამას, გარდა ტექსტის გახმოვანების შესაძლებლობისა, აქვს ორი სერვისი. პირველი თვალყურს ადევნებს ბაზის შემადგენლობას. თუ აღმოაჩენს, რომ ბაზაში არ გვაქვს რომელიმე ალოფონი, გვაძლევს შეტყობინებას ალოფონის მითითებით და ასრულებს მუშაობს. მეორე სერვისი სინთეზის პროცესში ითვლის და ბოლოს გვატყობინებს, თუ წაკითხული ტექსტის რამდენ პროცენტში მონაწილეობდა ძირითადი ბაზის ელემენტები და რამდენი იყო აღებული დამხმარე ბაზიდან.

ბაზის პერიოდული შევსების შემდეგ შესაძლებლობა გვაქვს მოვისმინოთ შედეგი და თუ საჭიროდ ჩავთვლით შევწყვიტოთ ბაზის შევსება. შეწყვეტის მიზეზი შეიძლება იყოს ორი, დამაკმაყოფილებელი ხარისხი ან პირიქით, მისი მიღების შესაძლებლობაში ეჭვის შეტანა. მეორე შემთხვევაში საჭიროა გავარკვიოთ მიზეზი და შევცვალოთ მსგავსების კლასები, ან გავზარდოთ მათი რიცხვი. შევცვალოთ წარმომადგენელთა არჩევის წესი, და სხვა, გამომდინარე დადგენილი მიზეზიდან.

ტესტირების მიმდინარეობისას გამოვლენილინდა ერთი ტენდენცია:

სინთეზის დროს მიიღებოდა მეტყველების გარკვეული არათანაბრობა, რის გამოც შევისწავლეთ ბგერის ხანგრძლივობის დამოკიდებულება სიტყვის სიგრძეზე. მივიღეთ საინტერესო და ლოგიკური შედეგი: ერთი და იმავე ბგერის საშუალო სიგრძე მნიშვნელოვნად განსხვავდება გრძელ და მოკლე სიტყვებში.



სურ.6. ბგერის ხანგრძლივობის დამოკიდებულების გრაფიკი სიტყვის სიგრძეზე.

დამუშავებული იყო 2352 სიტყვა სიგრძით 1-დან 31-მდე სიმბოლო. ორდინატა აღნიშნავს მილიწამებს, აბსცისა სიტყვის სიგრძეს სიმბოლოებში. ტექსტში არ იყო 20, 21, 22, 23, 25, 26, 27, 28, 29, 30 სიმბოლოს სიგრძის სიტყვები. მათი შესაბამისი უბანი გრაფიკზე ნულის ტოლია. ნახაზიდან კარგად ჩანს, რომ მათი არსებობის შემთხვევაში შესაბამისი წერტილები დაიკავებდნენ სრულიად განსაზღვრულ ადგილს წირზე. ეს მიუთითებს იმაზე, რომ სიტყვის სიგრძე და ბგერის ხანგრძლივობა ცალსახად არიან ერთმანეთზე დამოკიდებულნი. ჩვენს ბაზაში მოხვდნენ არათანაზომადი ერთეულები, რაც იწვევს მეტყველების არათანაზომობას.

ეს გრაფიკი უნდა ჩავთვალოთ, როგორც ერთი კონკრეტული დიქტორისათვის დამახასიათებელი. მისი განზოგადება, ვიდრე არ გვაქვს სხვა მონაცემები, არ შეიძლება. ტენდენცია ალბათ შენარჩუნდება. კონკრეტული მნიშვნელობები შესაძლებელია შეიცვალოს დიქტორიდან დიქტორამდე.

ამ მოვლენის კორექციის და კომპენსირების მიზნით გამოვთვალეთ ხმოვნის და თანხმოვნის საშუალო სიგრძეები, ისე რომ სინთეზის დროს მაქსიმალურად მივუახლოვდეთ სიტყვის რეალურ ხანგრძლივობებს. ეს არის ოპტიმიზაციის მარტივი ამოცანა.

აღვნიშნოთ:

L_v - ხმოვნის საშუალო სიგრძე;

L_c - თანხმოვნის საშუალო სიგრძე;

$\mathbf{x} = (L_v, L_c)'$, (' აღნიშნავს ტრანსპონირების ოპერაციას) ;

L_s - სიტყვის ხანგრძლივობათა ვექტორი-სვეტი (მოცემულია);

N_v - სიტყვაში ხმოვნების რაოდენობის ვექტორი-სვეტი (მოცემულია);

N_c - სიტყვაში თანხმოვნების რაოდენობის ვექტორი-სვეტი (მოცემულია);

$A = (N_v N_c)$ მატრიცა;

$J(\mathbf{x}) = \|A\mathbf{x} - L_s\|$ - ცდომილების ფუნქციონალი.

ჩვენი ამოცანაა ვიპოვნოთ \mathbf{x}^* , რომლისთვისაც ფუნქციონალი $J(\mathbf{x})$ აღწევს მინიმუმს.

სამიგბელი x^* არის ეგრეთწოდებული შეუღლებული სისტემის ამონახსნი:

$$A'Ay = A'Lv.$$

შეუღლებულ სისტემაში როცხვითი მონაცემების შეტანის და მისი ამონახსნის შედეგად მივიღეთ ხმოვნისა და თანხმოვნის ხანგრძლივობის ოპტიმალური მნიშვნელობები.

$$L_v = 120.6803 \text{ (მილიწამი);}$$

$$L_c = 105.7596 \text{ (მილიწამი).}$$

მეექვსე სურათზე წითელი მონაკვეთი აღნიშნავს ხმოვნის, ხოლო მწვანე თანხმოვნის ოპტიმალურ ხანგრძლივობას.

9. დასკვნები

2012 წლის იანვრის მდგომარეობით შევსებულია ზუსტ ალოფონთა ბაზის - **TriPhone 2.5435** პროცენტი, ხოლო წარმომადგენელთა ბაზის - **RepCL - 17.6921** პროცენტი. შესაბამისად სინთეზის ხარისხი არ არის დამაკმაყოფილებელი. აუცილებელია წარმომადგენელთა ბაზის სრული შევსება. ზუსტ ალოფონთა ბაზის სრული შევსება გამოიწვევს სინთეზის ხარისხის კიდევ მეტად გაუმჯობესებას.

როგორც აღვნიშნეთ, ტესტირების პროგრამა, გამომდინარე მისი დანიშნულებიდან, არ ახორციელებს ბგერათა სეგმენტების დამუშავებას და გვაძლევს სინთეზს "როგორიც არის".

ბაზის სრულად შევსების შემდეგ მოსალოდნელია გამოვლენილი იქნეს ორი ტიპის უზუსტობა:

1. ცალკეული ალოფონის არასწორი მარკირება;
2. მსგავსების კლასების „არაკორექტული“ განსაზღვრა.

მეორე ტიპი უზუსტობის გამოსწორება ხდება მარტივად. სინთეზის და ბაზის ფორმირების ყველა ალგორითმი და ფუნქცია (სწორედ ამ მიზნით) ცხადი სახით არის დამოკიდებული კლასების განსაზღვრულ პარამეტრებზე. ამიტომ ახალი კლასების საზღვრების მიწოდებით მივიღებთ მათ შესაბამის ბაზებს და სინთეზს.

ჩვენ მიერ ზემოთ აღწერილი ალოფონური ბაზის განსაზღვრის წესებისა და ბაზის შევსების ალგორითმის მიზანია სინთეზის პროცესში ყოველი სეგმენტისათვის აკუსტიკური მანძილის შემცირება წინა და მომდევნო სეგმენტებზე განთავსებულ ბგერებთან.

მიუხედავად ამისა, სასურველია სიტყვათა ძირითადი ტონის კონტურის ჩაწერა, რათა გამოვრიცხოთ ორი მნიშვნელოვნად განსხვავებული ტონის მქონე სეგმენტის მეზობლობით გამოწვეული არასწორი ინტონაციები სიტყვებში. საჭიროების შემთხვევაში უნდა მოხდეს მათი დაახლოება შესაბამისი მეთოდებით.

სრულფასოვანი და მაღალი ხარისხით ჩაწერილი ბაზა არის აუცილებელი, მაგრამ არასაკმარისი პირობა კარგი სინთეზისათვის, ამიტომ შემდგომ ეტაპზე საჭიროა არსებული ბაზის გამოყენებით მივიღოთ მაღალი ხარისხის, ბუნებრივი ჟღერადობის

მქონე მეტყველება. ეს ამოცანა მოითხოვს შესაბამისი მეთოდების გამოყენებას, რომლებსაც აქ აღარ განვიხილავთ²⁹.

ამრიგად, წარმოდგენილი პროექტის ფარგლებში კვლევით ამ დროისათვის მიღებული გვაქვს გამოყენებითი ღირებულების მქონე შემდეგი შედეგები:

1. მინიმალური ალოფონური ბაზის შემადგენლობა, ტრიფონთა ჩამონათვალი;
2. მინიმალური ბაზის შემცველი ტექსტი;
3. ტექსტის ყოველ სიტყვის შესაბამისი მაღალი ხარისხის აუდიო ჩანაწერი;
4. ანბანის სიმბოლოთა კონტექსტების რეალიზების სიხშირეთა მატრიცები;

პირველი ორი მათგანის გამოყენებით /გარდა მიმდინარე გამოყენებისა/, შესაძლებელია ჩავწეროთ სხვა ბაზებიც /სხვადასხვა ასაკის მამაკაცის, ქალის და ბავშვის/; მესამე შესაძლებელია გამოვიყენოთ სხვა ტიპის ბაზების შესაქმნელად /დიფონური, მარცვლოვანი, ეგრეთწოდებული მარტივი მარცვლებით, ნახევარმარცვლოვანი და სხვა/.

მეოთხე შედეგს აქვს უფრო ზოგადი მნიშვნელობა, იგი ახასიათებს არა კონკრეტულ დიქტორს, არამედ ქართულ ტექსტებს. ის საშუალებას გვაძლევს განვსაზღვროთ მსგავსების სხვა კლასები და ავირჩიოთ წარმომადგენლები.

აუდიო ჩანაწერები შესაძლებელია გამოვიყენოთ სიგნალთა აკუსტიკური მახასიათებლების შესასწავლად და სხვა ექსპერიმენტებისათვის.

ოთხივე ჩამოთვლილი შედეგი თან ერთვის ამ ანგარიშს კომპაქტ-დისკზე ჩაწერილი ფაილების სახით.

ამჟამად შესაძლებლობა გვაქვს მოვისმინოთ ნებისმიერი ტექსტის სინთეზი; თუმცა, ისეთი ტექსტების სინთეზი, რომლებიც შეიცავენ ისეთ ალოფონებს, რომელთა

²⁹მოგაწვდით ზოგიერთ სხვა ინფორმაციასაც, რაც, ჩვენი აზრით, ჩვენს სინამდვილეშიც გასათვალისწინებელია მომავალში: მეტყველების ბუნებრიობის მაღალი დონის თანამედროვე სინთეზატორებში ერთეულთა ბაზის მოცულობა შეადგენს ერთიდან ორ გიგაბაიტამდე. ჩვენს მიერ ამჟამად ჩაწერილი ბაზის მოცულობა შეადგენს 40 მეგაბაიტს. საშუალოდ 40-ჯერ ნაკლებია აღნიშნულზე. საჭიროა აგრეთვე სპეციფიკური ხასიათის გამოკვლევების წარმოება ქართული მეტყველების იმ პარამეტრების განსაზღვრისათვის, რომლებსაც ითხოვენ თანამედროვე მეტყველების და ამოცნობის ტექნოლოგიები; მაგალითისათვის, ინტელის დაკვეთით შეიქმნა მეტყველების კორპუსი რუსული ენისათვის [11], სადაც სამუშაოთა პირველ ეტაპზე შეიქმნა ხმის ჩაწერის, დამუშავების, სეგმენტაციის, ფილტრაციის, ვიზუალიზაციის და სხვა საჭირო პროგრამული უზრუნველყოფა. მეორე ეტაპზე, მათი გამოყენებით ჩაიწერა 237 დიქტორი (127 მამაკაცი, 110 ქალი) 18-დან 65 წლამდე, 6 ასაკობრივ კატეგორიაში. სხვა მონაცემებთან ერთად ფიქსირდებოდა მათი მისამართი (გეოგრაფიული მდებარეობა) განათლება და სპეციალობა. ჩანაწერის მოცულობამ შეადგინა 15 გიგაბაიტი (ჩვენს მიერ ჩაწერილია ერთი დიქტორი, ჩანაწერის მოცულობა შეადგენს 0,15 გიგაბაიტს).

ანალოგიური ტიპის და მოცულობის სამუშაოთა ჩატარება აუცილებლად მიგვაჩნია რეალურად ხელშესახები შედეგების მისაღწევად ქართული მეტყველების და ამოცნობის პროგრამირების ამოცანების გადაწყვეტის მიმართულებით.

აკუსტიკური ერთეულებიც ჯერ არ გვაქვს ბაზაში /ბაზის არასრული მარკირების გამო/, სინთეზის ხარისხი არ არის დამაკმაყოფილებელი.

სინთეზის ხარისხი, როდესაც ტექსტს ფარავს ბაზაში არსებული ალოფონები, მნიშვნელოვნად მაღალია და გამოირჩევა ნატურალური ჟღერადობით. ეს გვამღევეს იმედს, რომ სრულად მარკირებულ ბაზაზე აგებული სინთეზატორი მოგვცემს ჩვენთვის სასურველი, ნატურალური ჟღერადობის მქონე მეტყველებას.

ლიტერატურა

1. **Paul Taylor**. Text-to-Speech Synthesis. Cambridge University Press 2009, 642.
2. D. Jurafsky, J. H. Martin. Speech and Language Processing. Prentice-Hall Inc, 2000, 975.
3. Huedohg h., Acero A., Hsiao-vuen H. Spoken language processing. Prentice-Holl Inc, 2001, 935.
4. (<http://en.wikipedia.org/wiki/Triphone>).
5. В. С. Медведев, В. Г. Потемкин. Нейронные сети. MatLab 6, ДИАЛОГ-МИФИ, 2002, 496с.
6. А. Г. Фролов, Г. В. Фролов. Синтез и распознавание речи. Современные решения. 2003, 211с. <http://frolov-lib.ru/books/hi/ch01.html>.
7. В. Дьяконов, И. Абраменко. MATLAB. Обработка сигналов и изображений. Специальный справочник. ПИТЕР, 2002, 608с.
8. Ю. Кетков, А. Кетков, М. Шульц. MATLAB 7. Программирование, Численные методы. БВХ Петербург, 2005, 752с.
9. Л. М. Захаров. МГУ. Акустическая вариативность звуковых единиц в русской речи. 2004, 19с. <http://www.russian.slavica.org/article9312.html>.
10. Б.М. Лобанов, Л.И. Цирульник. ПРАВИЛА РАЗМЕТКИ РЕЧЕВОГО КОРПУСА НА ФОНЕТИЧЕСКИЕ СЕГМЕНТЫ И СТРАТЕГИЯ ВЫБОРА ЭЛЕМЕНТОВ КОМПИЛЯЦИИ ПРИ СИНТЕЗЕ РЕЧИ(НАН Беларуси, Минск, Беларусь).2007, 10с. <http://www.russian.slavica.org/article5578.html>.
11. Богданов Д. С., Кривнова О. Ф., Подрабинович А. Я. Современный инструментарии для разработки речевых технологии. Институт системного анализа РАН &МГУ. 22с. <http://www.philol.msu.ru/~otipl/SpeechGroup/publications.html>.

ალექსანდრე ვაშალომიძე,
ტექნიკურ. მეც. დოქტორი

როლანდი მათიაშვილი

ქართული ენის ტექნოლოგიების ცენტრის ვებ-გვერდის ბაზაზე პროექტის: „ნაბეჭდი ტექსტის ქართულად გახმოვანება და II საფეხურის ქართული სინტაქსური მართლმწერი“ გვერდის შექმნა.

ქართული ენის ტექნოლოგიების ცენტრის (მისამართი: <http://gtu.ge/gltc/index.htm/>) ძირითადი მიზანია ქართული ენის ტექნოლოგიება - ბუნებრივი ქართული ენისთვის ისეთი პროგრამების შექმნა, რომ თავისუფლად იყოს შესაძლებელი კომპიუტერთან (მანქანასთან) თავისუფალი ინტელექტუალური ურთიერთობა.

საქართველოს პოლიტექნიკური უნივერსიტეტის მიერ დაფინანსებული პროექტის /„ნაბეჭდი ტექსტის ქართულად გახმოვანება და II საფეხურის ქართული სინტაქსური მართლმწერი“/ მიზანია ამ მიმართულებით გარკვეული აქტივობების განხორციელება. წარმოდგენილი პროექტის ფარგლებში ჩვენი ამოცანა იყო ცენტრის ვებ გვერდის შექმნა და მასზე პროექტის მასალების განთავსება.

წარმოგიდგენთ საიტის სტრუქტურასა და შინაარსს:

ვებ-გვერდის ძირითადი პუნქტები და მენიუ:

- | | |
|-------------------------|-----------------------|
| 1. მთავარი გვერდი | 9. მონოგრაფიები |
| 2. ჩვენს შესახებ | 10. სტატიები. |
| 3. პროექტები | 11. თეზისები |
| 4. არქივი | 12. სახელმძღვანელოები |
| 5. სადემონსტრაციო კუთხე | 13. ბიბლიოთეკა |
| 6. ტექსტების კორპუსი | 14. სწავლება |
| 7. ფორუმი | 15. ვიდეო სიუჟეტები |
| 8. კონტაქტი | 16. პარტნიორები |

ქართული ენის ტექნოლოგიების ცენტრი Center For Georgian Language Technology

საქართველოს ტექნიკური უნივერსიტეტი



სიახლანი და მოვლენა

საქართველოს სახელმწიფო ენის დღისადმი მიძღვნილი კონფერენცია
სახელმწიფო ენა - ქვეყნის ღირსების საზოგადოებრივი კონფერენციის სამუშაო ენა - ქართული ენა

ყველა სიახლე

მთავარი გვერდი ჩვენს შესახებ ბიბლიოთეკა სადემონსტრაციო კუთხე ტექსტების კორპუსი ფორუმი კონტაქტი

მონოგრაფიები

- ქართული ენის II საფეხურის მათემატიკური თეორია
- ლოგიკური ბრუნებისა და ლინგვი-სტური მიმართებების სავითხისათვის ქართულში
- წინააღმდეგობრივი სემანტიკური ერთეულის ცნება, პირის ნიშანთა სავითხის

სტატიები

- ქართული ენობრივი ხედვა-ქართული ინტელექტუალური კომპიუტერული სისტემის აკვების ჩვენული გზა
- ქართული მკითხველი სისტემა სამო-მხმარებლო ხმის ჩამდები საშუალებით
- ქართული ენის თვისებები
- ხმით მართვადი ქართული ინტელექტუალური კომპიუტერული სისტემის I ვერსია

თვისებები

- ქართული ენის პირველი ნაწილი როგორც მათემატიკური თეორია
- ქართული ენის ლოგიკური გრამატიკა

სამი ღვთაებრივი საუნჯე დაგვჩაქართველებს მამა-პაპათან: „მამული, ენა, სარწმუნოება“ თუ ამასაც არ ეუპატრონეთ. რა ხალხი ვიქნებით, რა პასუხს გავცემთ შთამომავლობას.

ილია ჭავჭავაძე

მხოლოდ ის ერი შეიძლება დაადგეს პროგრესის გზას, რომელიც აზროვნებს და მეტყველებს მშობლიურ ენაზე.

დომინიკი ვიფიანი

მხოლოდ დედაენა ხსნის ყოველ ადამიანში ერთს განსაკუთრებულ ნიჭსა და ეს ნიჭი გახლავთ ალღო ენისა.

იაკობ გოგებაშვილი



ცნობილია, რომ ადამიანის ინტელექტუალური შესაძლებლობა მაქსიმალურად ვლინდება, თუ იმ აზროვნებს და მეტყველებს დედაენაზე. ის ერი (საზოგადოება), რომლის დედაენას არა აქვს კულტურულ-საგანმანათლებლო ფუნქცია, ან ჰქონდა, მაგრამ ვარგავს მას, არსებითი საფრთხის წინაშე დგას.

თანამედროვე ტექნოლოგიების სამყაროში, მხოლოდ ის ენები შეინარჩუნებენ კულტურული ენის სტატუსს, რომლებიც სრულფასოვნად ჩაერთვებიან გლობალურ ქსელში.

ქართველი ერის მიერ შექმნილი ენა - ქართველთა დედაენა - მრავალი საუკუნის მანძილზე უწყვეტად იყო საქართველოს სახელმწიფო ენა, საქართველოს სავლესიო ენა და ქართველი ერის საშვიგნობრო კულტურის ენა. დღეს ქართული ენა სახიცოცხლოდ მნიშვნელოვანი გამოწვევის წინაშეა, ვინაიდან, მსოფლიოს წამყვანი ენებისგან განსხვავებით, ქართული ენა უცერჯერობით სავმარისად არ არის ჩართული თანამედროვე ტექნოლოგიების ქსელში: დღემდე ვერ შეიქმნა ქართულად/ქართულიდან ავტომატურად მთარგმნელი კომპიუტერული სისტემა და სხვა ინტელექტუალური ქართულენოვანი კომპიუტერული სისტემები; შესაბამისად, დიდია ალბათობა, ქართულმა ენამ დაეარგოს კულტურულ-საგანმანათლებლო ენის ფუნქცია.

საქართველოს ტექნიკურ უნივერსიტეტთან არსებული "ქართული ენის ტექნოლოგიების სახელმწიფო-სამეცნიერო ცენტრის" მიზნებია: ბუნებრივი ენების ფორმალიზების დასავლური გამოყენების

პუნქტების შინაარსი და დანიშნულება:

- 1. მთავარი გვერდი** - წარმოაგენს საიტის Home Pages(ინდექს) და მასში თავმოყრილია საიტის ყველა პუნქტი და მენიუ.
- 2. ჩვენს შესახებ** – ასახულია მონაცემები ცენტრის თანამშრომელა შესახებ. ასვე წარდგენის პროექტი, პროექტის მიზნებისა და ძირითადი ამოცანების მოკლე ანოტაცია.
- 3. პროექტები**
 - ნაბეჭდი ტექსტის ქართულად გახმოვანება და II საფეხურის ქართული სინტაქსური მართლმწერი
 - ქართული ენის ლოგიკური გრამატიკის საფუძვლები და მისი გამოყენება საინფორმაციო ტექნოლოგიებში
- 4. არქივი** – წინა პერიოდის სტატიები და მოხსენებები
- 5. სადემონსტრაციო კუთხე** – პროგრამის მუშაობის ონლაინ ვერსია და აუდიო ფაილები
- 6. ტექსტების კორპუსი** – კორპუსითავისისტრუქტურითიძლევაქართულიენობრივისინამდვილის სრულყოფილადწარმოჩენისშესაძლებლობას.

ტექსტურ კორპუსში განთავსებული ლიტერატურა:

აკაკიწერეთელი
ალექსანდრეჭავჭავაძე
ჭაბუამირეჯიბი
ანაკალანდაძე
ბესიკი
სერგოკლდიაშვილი
დავითგურამიშვილი
გალაკტიონტაბიძე
გიორგილეონიძე
გოდერძიძე
გრიგოლორბელიანი
გრიგოლორბაქიძე
გურამდოჩანაშვილი
გურამრჩულიშვილი
იაკობგოგებაშვილი
იაკობხუცესი
ილიაჭავჭავაძე
იოანეზოსიძე
მიხეილჯავახიშვილი
კონსტანტინეგამსახურდია
მერაბკოსტავა
ნიკო ლომაური

ნოდარდუმბაძე
რაფიელერისთავი
რევაზინანიშვილი
შარლოდღერი
შექსპირი
შოთარუსთაველი
ტერენტიგრანელი
უიარალო
ვახტანგორბელიანი
ვაჟაფშაველა
ალიომირცხულავა
ალექსმენგელი
ვახტანგჯავახაძე
ვიქტორგაბისკირია
ზვიადგამსახურდია
მიხაბელაშვილი
სიმონჩიქოვანი
ნიკოსამადაშვილი
ნოდარჯალაღონია
პაოლოიაშვილი
ნიკოლოზბარათაშვილი
ნიკოლო ზლორთქიფანიძე

- 7. ფორუმი** – ფორუმზე გამოსატანი თემები შესათანხმებელია. ფორუმის ტექნიკური გამართვა დამოკიდებულია უნივერსიტეტის ბაზაზე არსებულ საშვალეებზე.

8. კონტაქტი – ცენტრის იურიდიული მისამართი და ხელმძღვანელის ელექტრო ფოსტა

9. მონოგრაფიები :

- ქართული ენის II საფეხურის მათემატიკური თეორია
- ლოგიკური ბრუნებისა და ლინგვისტური მიმართებების საკითხისათვის ქართულში
- წინარე ზმნური სემანტიკური ერთეულის ცნება, პირის ნიშანთა საკითხი
- ქართველოლოგიის ორგვარი დაგეგმვა – ტარიელ ფუტკარაძე, ია ვაშაკიძე
- სათავე და პერსპექტივა ქართული სამწიგნობრო ენისა
- "ევროპული ქარტია რეგიონული ან უმცირესობის ენის შესახებ"
- სათავე და პერსპექტივა ქართული სამწიგნობრო ენისა - ტ.ფუტკარაძე
- ქართულიენა - ტ.ფუტკარაძე
- ქართველები -ტ.ფუტკარაძე

10. სტატიები:

- ქართული ენობრივი ხედვა-ქართული ინტელექტუალური კომპიუტერული სისტემის აგების ჩვენეული გზა – Konstantine Pkhakadze, Aleksandre Maskharashvili, Lasha Abzianidze, Nikoloz Pkhakadze, Merab Chikvinidze
- ქართული მკითხველი სისტემა სამომხმარებლო ხმის ჩამდები საშუალებით – Pkhakadze K., Chichua G., Vashalomidze A., Abzianidze L., Maskharashvili A., Chikvinidze M.
- ხმით მართვადი ქართული ინტელექტუალური კომპიუტერული სისტემის I ვერსია – Pkhakadze K., Chichua G., Abzianidze L., Maskharashvil
- ქართული ენის თეზისები – Pkhakadze K., Maskharashvili A., Abzianidze L.
- დედაენა, სახელმწიფო ენა და თვითაღქმა
- თანამედროვე ქართველთა ენობრივ-ეთნიკური აღწერა "მსოფლიოს ენათა რეესტრის" მიხედვით და არსებული სინამდვილე
- სიტყვათშეხამება - ტ.ფუტკარაძე
- საქართველოს "ორგვარი" ლიგვისტურ-ეთნიკურ-პოლიტიკური ისტორია და ადეკვატური რუკების პრობლემა
- ქართველური დიალექტური ლექსიკა -სამწიგნობრო ქართულის მასაზრდოებელი წყარო
- სახელის ფორმამაწარმოებელი აფიქსების ისტორიული მიმართებისათვის ქართველური ქვესისტემების მონაცემების მიხედვით
- პალატალიზებული ხმოვნები ქართველურ ენა-კილოებში
- მოთხრობითი ბრუნვის ნიშნის ვარიანტთა ისტორიისათვის ქართული სამწიგნობრო ენისა და ქართველური დიალექტების მონაცემების მიხედვით
- ფორმა მაწარმოებელი ფორმანტების ისტორიული მიმართებისათვის
- ვითარებისა და მიმართულება-დანაშნულულების გამომხატავ ბრუნვის ნიშანთა ალომორფების შეპირისპირებითი ანალიზი
- ბუნთურქნი: პალეოგრაფიული შეცდომა
- პალატალიზებული ხმოვნები ქართველურ ენა-კილოებში და საქართველოს ისტორიის ზოგი საკითხი
- ოლთისი: სახელისეტიმოლოგიისადაცხე-ქალაქისრეალური ისტორიისათვის
- "ივერიელთა ერთობის ტრაქტატის" ტექსტის სემანტიკური ანალიზი
- ქართული ზმნის ლოგიკური პარადიგმის საკითხისათვის – ტ.ფუტკარაძე
- ისტორიული კოლხა და ქართველთა სამწიგნობრო ენა– ტ.ფუტკარაძე

11. თეზისები:

- ქართული ენის ძირეული ნაწილი როგორც მათემატიკური თეორია
- ქართული ენის ლოგიკური გრამატიკა და ქართულენოვანი კომპიუტერი – კონსტანტინე ფხაკაძე, ლაშა აბზიანიძე, ალექსანდრე მასხარაშვილი, ნიკოლოზ ფხაკაძე, მერაბ ჩიქვინიძე

- ქართული ენისა და აზროვნების მათემატიკური თეორიის სასაფუძვლო საკითხები და ხმით მართვადი ქართული კომპიუტერული სისტემის 1-ვერსიის ნაწილები
- ქართული ენის სრული მათემატიკური და მანქანური დაფუძნებისათვის

12. სახელმძღვანელოები

- გამოთვლითი ენათმეცნიერების საფუძვლები – ნაწილი II: გრამატიკული თეორიები
- მათემატიკური მეთოდები ენათმეცნიერებაში – ნაკვეთი I: მათემატიკური და ბუნებრივი ენების ზოგადი ლოგიკური საფუძვლები
- ლოგიკა და ენა: ტერმინთა მცირე განმარტებითი ლექსიკონი
- ქართული ენის ისტორია

13. ბიბლიოთეკა – ნაშრომები სხვადასხვა ენებზე :

- სამეცნიერო-საგანმანათლებლო ჟურნალი **ქართულიენადალოგიკა** №1 (იანვარი-ივნისი 2005წ.)
- Barbara H. Partee, Alice ter Meulen, Robert E. Wall **Mathematical Methods in Linguistics (თარგმანი) მათემატიკური მეთოდები ენათმეცნიერებაში** მათემატიკური და ბუნებრივი ენების ზოგადი ლოგიკური საფუძვლები
- ძირითადი და დამხმარე სახელმძღვანელოები თანამედროვე მათემატიკურ ენათმეცნიერებაში **R. Hausser - Foundations of Computational Linguistics (თარგმანი) გამოთვლითი ენათმეცნიერების საფუძვლები** ნაწილი II - გრამატიკული თეორიები
- **ქართული ენის ლოგიკა და სასწავლო უნარები** სამეცნიერო-საგანმანათლებლო ჟურნალის "ქართული ენა და ლოგიკა" ყოველ წლიური დამატება №1 (2007 წ.)მესამე დია რესპუბლიკური სასემინარო კვირეული **ლოგიკა, ენა, გამომთვლელი** სასემინარო კვირეულის მოხსენებები (11-18 აპრილი 2007 წ.)
- სამეცნიერო-საგანმანათლებლო ჟურნალი **ქართული ენა და ლოგიკა** სპეციალურიგამოცემა №3, №4, №5, №6 - I ნაწილი (2006-2007 წ.)
- ტარილ ფუტკარაძე - **საქართველოს ისტორია** (მოკლევარიანტი)
- უწმინდესისა და უნეტარესის, სრულიად საქართველოს კათოლიკოს-პატრიარქის **ილია II-ის** საკვირაო ქადაგება
- ენისა და დიალექტის საკითხი ქართველოლოგიაში – თეიმურაზგვანცელაძე
- ივანე ჯავახიშვილისა და ნიკო მარის მიმოწერა

14. სწავლება

- საბაკალავრო სწავლება
- სამაგისტრო სწავლება
- სადოქტორო სწავლება

15. ვიდეო სიუჟეტები – ქართული ინტელექტუალური კომპიუტერული სისტემები

- პირველი არხი
- რუსთავი-2
- საზღვარგარეთშექმნილიტექსტის დახმისამომცნობიინტელექტუალურიპროგრამები
- ცენტრის მუშაობის ამსახველი ფილმი – მიმდინარეობს ფილმის ვიდეო მონტაჟი

16. პარტნიორები

საიტის დიზაინზედა მაკეტზე მუშაობის პროცესი

ebi\New Folder (2)\html-1\index.html

GLTC საქართველოს ტექნიკური უნივერსიტეტი GLTC GLTC New Page 1 GLT X GLTC

GLTC Georgian Language Technology Center

ქართული ენის ცოდნოლოგიის ცენტრი

თავფურცელი ცენტრის შესახებ პუბლიკაციები სიახლეები ბიბლიოთეკა არქივი ფორუმი კონტაქტი



www.GLTC.ge

ინფორმაცია

- სახელმწიფო ენის პროგრამა
- ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრი
- ქართული ენის ლოგო-სასწავლო უზარები
- Toward Logical Grammar of Georgian language and Georgian Computer
- ინფორმატიკის მიმართულება
- პატრიარქის ქადაგება
- ქართული ენის ლოგიკური გრამატიკა...

ლაბორატორია

- ხმოვანი ბაზის სინთეზის წესები
- ქართული ტექსტის სტატისტიკური ანალიზისათვის
- „ქართული ენის ლოგიკური გრამატიკა და ქართულენოვანი კომპიუტერი“
- ბებ/ და /ზეზ/ შედარება
- მიკროკომპიუტერის სინთეზის

„ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის“ საქმიანობის მიზნები, ძირითადი მიმართულებები და პირველ რიგში გადასაჭრელი ამოცანები

2010 წელს საქართველოს ტექნიკურ უნივერსიტეტში დაფუძნდა ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრი. ცენტრი ამჟამად ხანგრძლივ ვადიან პროექტს „ქართული ენის ლოგიკური გრამატიკა და ქართულენოვანი კომპიუტერი“. პროექტი, რომელიც მანამდე არსებული სახელმწიფო-მიზნობრივი პროგრამის „კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ კვლევითი ამოცანების შემდგომი განვითარების საფუძველზე ჩამოყალიბდა, მიზნად ისახავს ქართული ენის ბუნების სრულად ამსახველი მათემატიკური თეორიის აწლ, მოკლედ, ქართული ენის ლოგიკური გრამატიკის შემუშავებას და, ასევე, ქართულენოვანი კომპიუტერის აწლ ინტელექტუალური უზარებით აღჭურვილი ქართული კომპიუტერული სისტემის აგებას.

ხაზგასმელია, რომ ქართული ინტელექტუალური კომპიუტერული სისტემის აგების ერთდერთი გზა ქართულ ენით მოცემული და, ამდენად, ქართულ ენაში არსებული ინტელექტუალური გამოთვლების სრულ მოდელირებაზე გადის, რაც ქართული ენის სრული მათემატიკური თეორიის შემუშავებას გულისხმობს – აქ, ქართული ენის ქვეშ გაიგება ქართული ბუნებრივი ენობრივი სისტემა, რომლის შემადგენლებიცაა ქართული სამწერლობო, სამეტყველო და სააზროვნო ენები. ამგვარად, მოხსენებისას მიმოვიხილავთ ქართული ენის მათემატიკური თეორიის აწლ ქართული ენის ლოგიკური გრამატიკის საფუძველმდებით საკითხებს, ესენია:

- ლოგიკური ბრუნების და ლინგვისტური პრედიკატის ცნება, რომელთა საფუძველზე შემუშავებულია ახალი მათემატიკური ხედვა ქართული ენის არსებითი და ზედსართავი სახელებისა და ზმნების მორფოლოგიურ, სინტაქსურ, ლოგიკურ და სემანტიკურ ბუნებაზე. ასევე, ხაზგასმული იქნება ქართულ და ინდო-ევროპულ ნაცვალსახელებს, ზმნებსა და უღლებით სისტემებს შორის არსებული პრინციპული განსხვავებები;
- ქართული ენის I საფეხურის მათემატიკური თეორია, რომელიც სრულად აღწერს ქართული ენის ძირითად ნაწილს აწლ ქართული ენის ფორმული ნაწილისათვის შრის აწლ ნარმოციანი ენის

Articles - Google ... ვა-დაზარის სტუ... C:\Documents an... D:\gpi kvelafero... E:\kartuli enis tek... GLT - Windows In... 4 Microsoft Offi...

II -ვერსია

GLTC საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრი
Scientific-Educational Center For Georgian Language Technology
At The Georgian Technical University

მთავარი
სიახლეები
არქივი
საუწყებო გვერდი
პუბლიკაციები
ბიბლიოთეკა
ცენტრის შესახებ
ცენტრის წევრები
პროექტები
ფორუმი
კონტაქტი

პირველი ვერსია

საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრი

2010 წელს საქართველოს ტექნიკურ უნივერსიტეტში დაფუძნდა ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრი. ცენტრი ამუშავებს ხანგრძლივ ვადიან პროექტს „ქართული ენის ლოგიკური გრამატიკა და ქართულენოვანი კომპიუტერი“. პროექტი, რომელიც მანამდე არსებული სახელმწიფო-მიზნობრივი პროგრამის „კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ კვლევითი ამოცანების შემდგომი განვითარების საფუძველზე ჩამოყალიბდა, მიზნად ისახავს ქართული ენის ბუნების სრულად ამსახველი მათემატიკური თეორიის ანუ, მოკლედ, ქართული ენის ლოგიკური გრამატიკის შემუშავებას და, ასევე, ქართულენოვანი კომპიუტერის ანუ ინტელექტუალური უზარებით აღჭურვილი ქართული კომპიუტერული სისტემის აგებას.

translation, arabic, machine, alignment, smt, statistical, output, dependency, human, experiments, also, data, show, sentence, framework, used, evaluation, results, english, investigate, system, bleu, smt, ქართული ენის ტექნოლოგიები

III -ვერსია

საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრი
Scientific-Educational Center For Georgian Language Technology At The Georgian Technical University

მთავარი ცენტრის შესახებ ბიბლიოთეკა ფორუმი კონტაქტი

GLTC

ახალი პუბლიკაციები

-
-
-
-
-

მიმდინარე მოვლენები

-
-
-
-

პროექტები

-

სამი ღვათებრივი საუნჯე დაგვრჩა ქართველებს მამა-პაპათგან: „მამული, ენა, სარწმუნოება“, თუ ამასაც არ ვუპატრონეთ. რა ხალხი ვიქნებით, რა პასუხს გავცემთ შთამომავლობას.
ილია ჭავჭავაძე

მხოლოდ ის ერი შეიძლება დაადგეს პროგრესის გზას, რომელიც აზროვნებს და მეტყველებს მშობლიურ ენაზე.
დიმიტრი ყიფიანი

მხოლოდ დედაენა ხსნის ყოველ ადამიანში ერთს განსაკუთრებულ ნიჭსა და ეს ნიჭი გახლავთ ალღო ენისა.
იაკობ გოგებაშვილი

ცნობილია, რომ ადამიანის ინტელექტუალური შესაძლებლობა მაქსიმალურად ვლინდება, თუკი ის აზროვნებს და მეტყველებს დედაენაზე. ის ერი (საზოგადოება), რომლის დედაენას არა აქვს კულტურულ – საგანმანათლებლო ფუნქცია, ან ჰქონდა, მაგრამ კარგავს მას, არსებითი საფრთხის წინაშე დგას. თანამედროვე ტექნოლოგიების სამყაროში, მხოლოდ ის ენები შეინარჩუნებენ კულტურული ენის სტატუსს, რომლებიც სრულფასოვნად ჩაერთვებიან გლობალურ ქსელში.

ქართველი ერის მიერ შექმნილი ენა - ქართველთა დედაენა - მრავალი საუკუნის მანძილზე უწყვეტად იყო საქართველოს სახელმწიფო ენა, საქართველოს საეკლესიო ენა და ქართველი ერის სამწიგნობრო კულტურის ენა. დღეს ქართული ენა სასიცოცხლოდ მნიშვნელოვანი გამოწვევის წინაშეა, ვინაიდან, მსოფლიოს წამყვანი ენებისგან განსხვავებით, ქართული ენა ჯერჯერობით საკმარისად არ არის ჩართული თანამედროვე ტექნოლოგიების ქსელში: დღემდე ვერ შეიქმნა ქართულად/ქართულიდან ავტომატურად მთარგმნელი კომპიუტერული სისტემა და სხვა ინტელექტუალური ქართულენოვანი კომპიუტერული სისტემები; შესაბამისად, დიდია ალბათობა, ქართულმა ენამ დაკარგოს კულტურულ-საგანმანათლებლო ენის ფუნქცია.

საქართველოს ტექნიკურ უნივერსიტეტთან არსებული "ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის" მიზნებია: ბუნებრივი ენების ფორმალიზების დასავლური ამოცილების აპირაციის და ავტომატური ამოცილების სამიზნო პროექტის განხორციელება.

IV -ვერსია

ersiebl\New Folder (2)\1 variant\index.htm

TC X GLTC საქართველოს ტექნიკ... GLTC GLTC New Page 1 GLT GLTC

English

ქართული ენის ტექნოლოგიების ცენტრი Georgian Language Technology Center

საქართველოს ტექნიკური უნივერსიტეტი



GLTC
თბილისი, კოსტავას 77
ადმინისტრაციული კორპუსი

სისალაპი და მოვლაპი
საქართველოს სახელმწიფო ენის დღისადმი მიძღვნილი კონფერენცია
სახელმწიფო ენა - ქვეყნის დირსების საზომი კონფერენციის სამუშაო ენა _ ქართული ენა
ველა სიახლე

მთავარი გვერდი ცენტრის შესახებ ბიბლიოთეკა პროექტები ფორუმი კონტაქტი

მიმდინარე მოვლენები

- 14 აპრილი - საქართველოს სახელმწიფო ენის დღე
- საქართველოს მეცნიერებათა ეროვნული აკადემიის ფიზიკა - მათემატიკის სექციის სემინარი
- ლინგვისტურ სწავლებათა დაფუძნების აღცილებლობა
- ინფორმაციული ტექნოლოგიები და კომპიუტერული მოდელირება

ახალი პუბლიკაციები

- ქართული ენის ლოგიკური გრამატიკა და ქართულენოვანი კომპიუტერი
- ქართული ინტელექტუალური კომპიუტერული სისტემის შემუშავების მიზნები და ქართული ენის წინაშე

სამი დვთაბერიფი საუნჯე დაგვრჩა ქართველებს მამა-პაპათან: „მამული, ენა, სარწმუნოება“ თუ ამასაც არ ვუპატრონეთ, რა ხალხი ვიქნებით, რა პასუხს გავცემთ შთამომავლობას.
ილია ჭავჭავაძე

მხოლოდ ის ერი შეიძლება დაადგეს პროგრესის გზას, რომელიც აზროვნებს და მეტყველებს მშობლიურ ენაზე.
დიმიტრი ყიფიანი

მხოლოდ დედაენა ხსნის ყოველ ადამიანში ერთს განსაკუთრებულ ნიჭსა და ეს ნიჭი გახლავთ ალღო ენისა.
იაკობ გოგებაშვილი



ცნობილია, რომ ადამიანის ინტელექტუალური შესაძლებლობა მაქსიმალურად ვლინდება, თუკი ის აზროვნებს და მეტყველებს დედაენაზე. ის ერი საზოგადოება), რომლის დედაენას არა აქვს კულტურულ - საგანმანათლებლო ფუნქცია, ან ჰქონდა, მაგრამ კარგავს მას, **არსებითი საფრთხის წინაშე დგას.** თანამედროვე ტექნოლოგიების სამყაროში, მხოლოდ ის ენები შეინარჩუნებენ კულტურული ენის სტატუსს, რომლებიც სრულფასოვნად ჩაერთვებიან გლობალურ ქსელში.
ქართველი ერის მიერ შექმნილი ენა - **ქართველთა დედაენა** - მრავალი საუკუნის მანძილზე უწყვეტად იყო

Articles - Google ... ვებ-დოკუმენტის სტუ... C:\Documents an... D:\gpi kvelafero... E:\kartuli enis tek... GLTC - Windows I... 4 Microsoft Offi...

ქართული ენის ტექნოლოგიების ცენტრი Center For Georgian Language Technology

საქართველოს ტექნიკური უნივერსიტეტი



სიახლეები და მოვლახეობა

საქართველოს სახელმწიფო ენის დღისადმი მიძღვნილი კონფერენცია
სახელმწიფო ენა - ქვეყნის ღირსების საზომი კონფერენციის სამუშაო ენა _ ქართული ენა

ყველა სიახლე

მთავარი გვერდი

ცენტრის შესახებ

ბიბლიოთეკა

პროექტები

ფორუმი

კონტაქტი

სტატიები

-
-
-
-

სახელმძღვანელოები

-
-
-
-

სწავლება

სამი ღვთაებრივი საუნჯე დაგვჩა ქართველებს მამა-პაპათანა: „მამული, ენა, სარწმუნოება“ თუ ამასაც არ ვუპატრონეთ. რა ხალხი ვიქნებით, რა პასუხს გავცემთ შთამომავლობას.

ილია ჭავჭავაძე

მხოლოდ ის ერი შეიძლება დაადგეს პროგრესის გზას, რომელიც აზროვნებს და მეტყველებს მშობლიურ ენაზე.

დიმიტრი ყიფიანი

მხოლოდ დედაენა ხსნის ყოველ ადამიანში ერთს განსაკუთრებულ ნიჭსა და ეს ნიჭი გახლავთ ალღო ენისა.

იაკობ გოგებაშვილი



ცნობილია, რომ ადამიანის ინტელექტუალური შესაძლებლობა მაქსიმალურად ვლინდება, თუკი ის აზროვნებს და მეტყველებს დედაენაზე. ის ერი (საზოგადოება), რომლის დედაენას არა აქვს კულტურულ – საგანმანათლებლო ფუნქცია, ან ჰქონდა, მაგრამ კარგავს მას, არსებითი საფრთხის წინაშე დგას. თანამედროვე ტექნოლოგიების სამყაროში, მხოლოდ ის ენები შეინარჩუნებენ კულტურული ენის სტატუსს, რომლებიც სრულფასოვნად ჩაერთვებიან გლობალურ ქსელში.

ქართული ერის მიერ შექმნილი ენა - ქართველთა დედაენა - მრავალი საუკუნის მანძილზე უწყვეტად იყო

VI -ვერსია

New Page 1 - Windows Internet Explorer
 D:\gpi kvelafero\gpiis saitis dzveli versiebi\New Folder (2)\4 variant\axali -1\index2.htm

File Edit View Favorites Tools Help
 Преобразовать Выбрать

CFGLT GLTC საქართველოს ტექნიკური უნივერსიტეტი GLTC GLTC New Page 1

საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრი

GLTC

Scientific-Educational Center For Georgian Language Technology At The Georgian Technical University

მთავარი ცენტრის შესახებ ბიბლიოთეკა ფორუმი კონტაქტი

GLTC

ახალი პუბლიკაციები

-
-
-
-
-

მიმდინარე მოვლენები

-
-
-
-
-

პროექტები

სამი ღვთაებრივი საუნჯე დაგვრჩა ქართველებს მამა-პაპათგან: „მამული, ენა, სარწმუნოება“, თუ ამასაც არ ვუპატრონეთ. რა ხალხი ვიქნებით, რა პასუხს გავცემთ შთამომავლობას.

ილია ჭავჭავაძე

მხოლოდ ის ერი შეიძლება დაადგეს პროგრესის გზას, რომელიც აზროვნებს და მეტყველებს მშობლიურ ენაზე

დიმიტრი ყიფიანი

მხოლოდ დედაენა ხსნის ყოველ ადამიანში ერთს განსაკუთრებულ ნიჭსა და ეს ნიჭი გახლავთ ალღო ენისა.

იაკობ გოგებაშვილი

ცნობილია, რომ ადამიანის ინტელექტუალური შესაძლებლობა მაქსიმალურად ვლინდება, თუკი ის აზროვნებს და მეტყველებს დედაენაზე. ის ერი (საზოგადოება), რომლის დედაენას არა აქვს კულტურულ –საგანმანათლებლო ფუნქცია, ან ჰქონდა, მაგრამ კარგავს მას, **არსებითი საფრთხის წინაშე დგას**. თანამედროვე ტექნოლოგიების სამყაროში, მხოლოდ ის ენები შეინარჩუნებენ კულტურული ენის სტატუსს, რომლებიც სრულფასოვნად ჩაერთებიან გლობალურ ქსელში.

ქართველი ერის მიერ შექმნილი ენა - **ქართველია დედაენა** - მრავალი საუკუნის მანძილზე უწყვეტად იყო საქართველოს სახელმწიფო ენა, საქართველოს საეკლესიო ენა და ქართველი ერის სამწიგნობრო კულტურის ენა. **დღეს ქართული ენა სასიცოცხლოდ მნიშვნელოვანი გამოწვევის წინაშეა**, ვინაიდან, მსოფლიოს წამყვანი ენებისგან განსხვავებით, ქართული ენა ჯერჯერობით საკმარისად არ არის ჩართული თანამედროვე ტექნოლოგიების ქსელში: **დღემდე ვერ შეიქმნა ქართულად/ქართულიდან ავტომატურად მთარგმნელი კომპიუტერული სისტემა და სხვა ინტელექტუალური ქართულენოვანი კომპიუტერული სისტემები**; შესაბამისად, დიდია ალბათობა, ქართულმა ენამ დაკარგოს კულტურულ-საგანმანათლებლო ენის ფუნქცია.

Done

ვებ საიტის მიზნებ... Articles - Google ... ვებ-დოკუმენტის სტ... C:\Documents an... D:\gpi kvelafero\... E:\kartuli enis tek... New

ქართული ენის ტექნოლოგიების ცენტრი Center For Georgian Language Technology

საქართველოს ტექნიკური უნივერსიტეტი



CFGLT
თბილისი, კოსტავას 77
ადმინისტრაციული კორპუსი

სისხლავა და მოვლანა

საქართველოს სახელმწიფოს დღისადმი მიძღვნილი კონფერენცია
სახელმწიფო - ქვეყნის ღირსების საზოგადოებრივი კონფერენციის სამუშაო ანგარიშის კარგადი

ყველა სიახლე

- მთავარი გვერდი
- ჩვენს შესახებ
- ბიბლიოთეკა
- სადემონსტრაციო კუთხე
- ტექსტების კორპუსი
- ფორუმი
- კონტაქტი

მონოგრაფიები

- ქართული ენის II საფეხურის მათემატიკური თეორია
- ლოგიკური ბრუნვისა და ლინგვი - სტური მიმართულების საკითხისათვის ქართულში
- წინარე ზმნური სემანტიკური ერთე - ულის ცნება, პირის ნიშანთა საკითხი

სტატიები

- ქართული ენობრივი ხედეა-ქართული ინტელექტუალური კომპიუტერული სისტემის აგების ჩვენური გზა
- ქართული შვითხევილი სისტემა სამო - მხმარებლო ხმის ჩამდები საშუალებით
- ქართული ენის თეზისები
- ხმით მართვადი ქართული ინტელექტუალური კომპიუტერული სისტემის I ვერსია

თეზისები

- ქართული ენის მირული ნაწილი როგორც მათემატიკური თეორია
- ქართული ენის ლოგიკური გრამატიკა

სამი დღეაბრივი სახეჯე დაგვრჩა ქართველებს მამა-პაპათან: „მამული, ენა, სარწმუნოება“ თუ ამასე არ ვუპატრონეთ, რა ხალხი ვიქნებით, რა პასუხს გავცემთ შთამომავლობას.

მხოლოდ ის ერი შეიძლება დაადგეს პროგრესის გზას, რომელიც აზროვნებს და მეტყველებს მშობლიურ ენაზე.

მხოლოდ დედაენა ხსნის ყოველ ადამიანში ერთს განსაკუთრებულ ნიჭსა და ეს ნიჭი განლავეთ ალღო ენისა.

იაკობ გოგებაშვილი



ენობილია, რომ ადამიანის ინტელექტუალური შესაძლებლობა მქსიმალურად ვლინდება, თუვი ის აზროვნებს და მეტყველებს დედაენაზე. ის ერი საზოგადოება, რომლის დედაენას არა აქვს ვულტურულ-საგანმანათლებლო ფუნქცია, ან ჰქონდა, მაგრამ ვარგავს მას, **არსებითი საფრთხის წინაშე დგას.**

თანამედროვე ტექნოლოგიების სამყაროში, მხოლოდ ის ენები შეინარჩუნებენ ვულტურული ენის სტატუსს, რომლებიც სრულფასოვნად ჩაერთვებიან გლობალურ ქსელში.

ქართველი ერის მიერ შექმნილი ენა - **ქართველი დედაენა** - მრავალი საუკუნის მანძილზე უწყვეტად იყო საქართველოს სახელმწიფო ენა, საქართველოს საეკლესიო ენა და ქართველი ერის საწინგნობრო ვულტურის ენა. **დღეს ქართული ენა სასიცოცხლოდ მნიშვნელოვანი გამოწვევის წინაშეა**, ვინაიდან, მსოფლიოს წამყვანი ენებისგან განსხვავებით, ქართული ენა ჯერჯერობით საგმარისად არ არის ჩართული თანამედროვე ტექნოლოგიების ქსელში: **დღემდე ვერ შეიქმნა ქართულად/ქართულიდან ავტომატურად მთარგმნელი კომპიუტერული სისტემა და სხვა ინტელექტუალური ქართულენოვანი კომპიუტერული სისტემები**; შესაბამისად, დიდია ალბათობა, ქართულმა ენამ დაეარგოს ვულტურულ-საგანმანათლებლო ენის ფუნქცია.

საქართველოს ტექნიკურ უნივერსიტეტთან არსებული **„ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრის“** მიზნებია: ზონიბრივი ენების დორმალიზების დასაქლორი გამოყადილოების