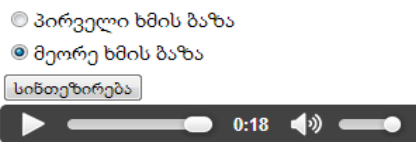
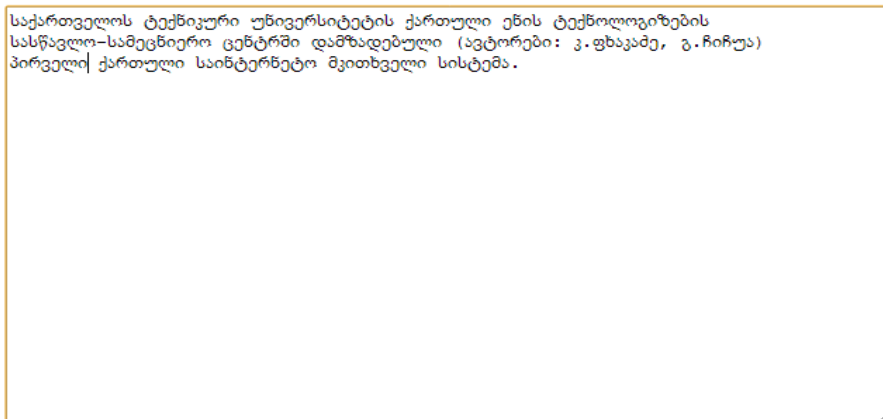


საქართველოს ტექნიკური უნივერსიტეტის გამოყენებითი საგრანტო
ნომინაციის №048-13 საპროექტო წინადადების
„ქართული ენის ტექნოლოგიური ანბანის ასაგებად აუცილებელი რიგი
სისტემების გაფართოებადი (სწავლებადი) საინტერნეტო ვერსიების
შემუშავება“

ანგარიში

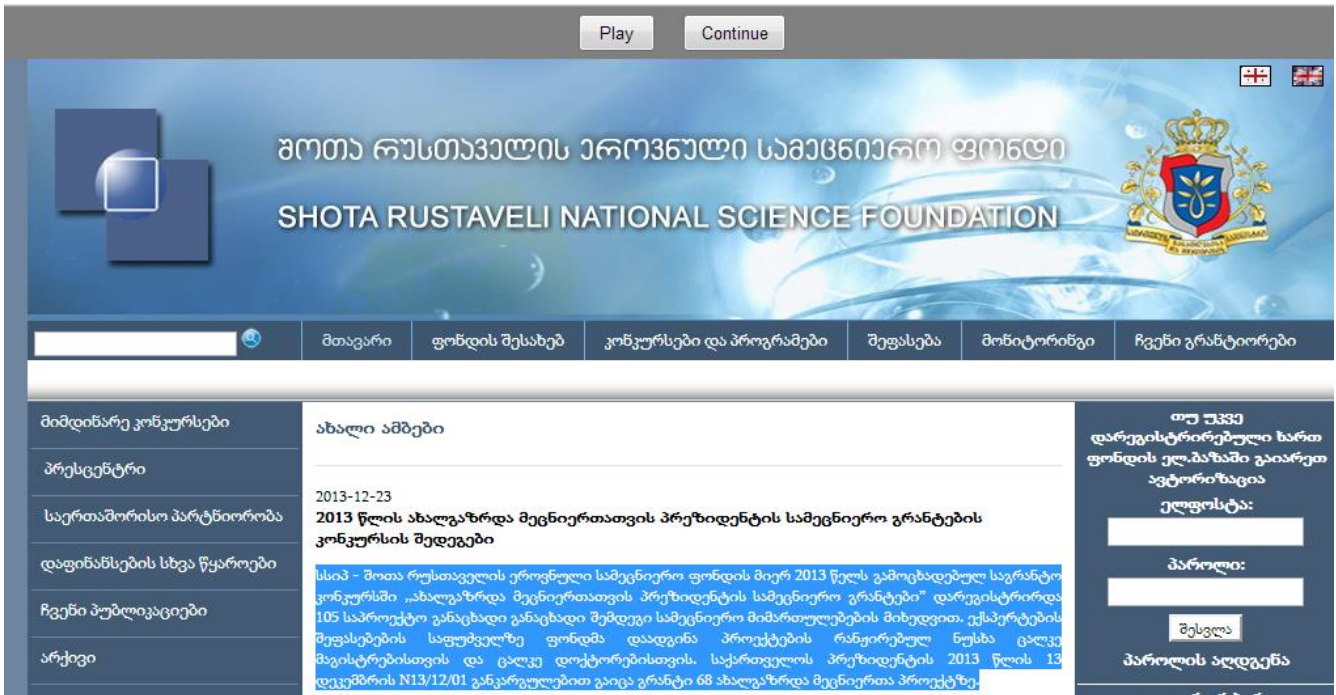
(2013.09.01 - 2014.03.31)

1. ხელშეკრულების მიხედვით ამოცანა N1-ით დაგეგმილი იყო ქართული ელ-ტექსტების მკითხველი საინტერნეტო სისტემის (განთავსებულია <http://www.geotextreader.com/> და <http://voice.dmind.net/> მისამართებზე) კითხვის ხარისხის თვისობრივი გაუმჯობესება. - დასახული ამოცანა შესრულდა სრულად. უფრო მეტიც, ამ ამოცანის გასწვრივ ჯგუფი გასცდა საპროექტო დავალებას და გარდა უკვე არსებული ქართული ელ-ტექსტების მკითხველი სისტემისა (ღია ფორმით განთავსებულია <http://geoanbani.com/> მისამართზე) შექმნა ქართული საიტების პირველი და ჯერ-ჯერობით ერთადერთი მკითხველი სისტემა.



სურათი 1

სურათ 1-ზე გამოსახულია საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების ცენტრში დამზადებული (ავტორები: კ.ფხაკაძე, გ.ჩიჩუა) პირველი ქართული საინტერნეტო მკითხველი სისტემა.



სურათი 2

მე-2-ე სურათზე გამოსახულია საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიზების ცენტრში დამზადებული (ავტორები: კ.ფხაკაძე, გ.ჩიჩუა, მ.ჩიქვინიძე) ვებ-გვერდების მკითხველი პირველი და ჯერ-ჯერობით ერთადერთი ქართული საინტერნეტო სისტემა მოქმედებაში, ანუ, მაშინ, როცა ის კითხულობს შოთა რუსთაველის ეროვნული სამეცნიერო ფონდის ვებ-გვერდს.

ამასთან, ხაზგასასმელია, და უფრო მეტიც, საგანგაშოა, რომ, სამწუხაროდ, ამ ჯერ-ჯერობით ერთადერთი ქართული ვებ-გვერდების მკითხველი სისტემის კითხვის ხარისხი ჯერ კიდევ არასაკმარისად მაღალია. თუმცა, თავიც რომ დავანებოთ „ქართული ენის ტექნოლოგიზების სასწავლო-სამეცნიერო ლაბორატორიას“, რომლის ფორმირება, სამწუხაროდ, ჯერ კიდევ ვერ მოხერხდა და რომლის ფორმირების შემდეგ ჩვენ აქ შედეგების შემდგომი სწრაფი გაუმჯობესების შანსები გვეძლევა, ხაზგასასმელია ისიც, რომ გამოთვლითი ლინგვისტიკის ფარგლებში ქართული მეტყველება მკითხველი სისტემების ასაგებად აუცილებელი თვალსაზრისებით არასაკმარისადაა შესწავლილი, რაც ვითარების ზემოაღნიშნული საგანგაშობის ხარისხს კიდევ უფრო მაღალს ხდის.

ამოცანა N1-ით განსაზღვრული საპროექტო დავალების შესრულებასთან დაკავშირებული პრობლემები: იმან, რომ შეფერხდა „ქართული ენის ტექნოლოგიზების სასწავლო-სამეცნიერო ლაბორატორიის“ ფორმირება, ხელი შეგვიშალა უკვე შემენილი ხმის ჩამწერი კომპლექტის სალაბორატორიო განთავსებაში, რამაც, თავის მხრივ, ხელი შეგვიშალა უფრო ხარისხიანი სახმო-სამეტყველო ბაზების შექმნასა და კომპიუტერული კითხვის ახალი ალგორითმის რეალიზაციაში.

2. ხელშეკრულების მიხედვით ამოცანა N2-ით დაგეგმილი იყო გაფართოებადი ქართული გრამატიკული მართლმწერის საინტერნეტო სისტემის (განთავსებულია <http://translate.dmind.net/Login.aspx> მისამართზე) გაფართოება და პირობითად

www.geogrammchecker.com მისამართზე განთავსება. - ამ ამოცანით დაგეგმილი სამუშაოები განხორციელდა სრულად. უფრო მეტიც, ჯგუფი ამ ამოცანის შემთხვევაშიც გასცდა საპროექტო დავალებას და უკვე ააგო პირველი და ჯერ-ჯერობით ერთადერთი ავტომატურად გაფართოებადი ქართული სინტაქსური მართლმწერის საინტერნეტო სისტემა (განთავსებულია <http://geoanbani.com/> მისამართზე.

კაცს დახატა ქალი. კაცმა ქალი დახატა. ქალმა კაცს ხატავს. ქალი კაცს ხატავს. თუ ქალი ხატავს კაცს, მაშინ კაცს უყვარს ქალი. თუ ქალი ხატავს კაცს მაშინ კაცს უყვარს ქალი. თუ ქალი კაცს ხატავს, მაშინ უყვარს. თუ ქალი უყვარს კაცს და ხატავს, ქალი წერს ლექსს. თუ ქალი უყვარს კაცს და ხატავს, მაშინ ქალი წერს ლექსს. ქალი, რომელიც უყვარს კაცს, ლექსს წერს. ქალი რომელიც უყვარს კაცს, ლექსს წერს. ქალმა, რომელიც უყვარდა კაცს, ლექსი დაწერა. ქალს, რომელიც უყვარდა კაცს, ლექსი დაწერა.



სურათი 3

მე-3-ე სურათზე გამოსახულია საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების ცენტრში დამზადებული (ავტორები: კ.ფხაკაძე, მ.ჩიქვინიძე) პირველი და ჯერ-ჯერობით ერთადერთი ქართული რთული და მარტივი წინადადებების ავტომატურად გაფართოებადი სინტაქსური მართლმწერი მოქმედებაში. კერძოდ, როგორც ეს სურათიდანაც ჩანს, სისტემამ გამოიცნო და ყვითლად შეაფერადა სინტაქსურად გაუმართავი წინადადებები. ამასთან, გაყვითლებულ ადგილზე კურსორის მიტანით სისტემა მომხმარებელს ეკრანზე გამოსული „ფანჯრის“ მეშვეობით აწვდის ინფორმაციას გაყვითლების მიზეზის ანუ გაყვითლებულ ფრაზაში დაშვებული შეცდომის თაობაზე.

ხაზგასასმელია, და უფრო მეტიც, საგანგაშოა, რომ, სამწუხაროდ, ამ ჯერ-ჯერობით ერთადერთი ქართული გრამატიკული ანუ სინტაქსური მართლმწერის მოქმედების ხარისხი და არეალი ჯერ კიდევ ძლიერ შეზღუდულია. თუმცა, ხაზგასასმელია ისიც, რომ გამოთვლითი ანუ მათემატიკური ლინგვისტიკის ფარგლებში ქართული ენა მართლმწერი სისტემის ასაგებად აუცილებელი თვალსაზრისებით სრულიად არასაკმარისად არის შესწავლილი, რაც ვითარების ზემოაღნიშნული საგანგაშობის ხარისხს კიდევ უფრო მაღალს ხდის.

ამოცანა N2-ით განსაზღვრული საპროექტო დავალების შესრულებასთან დაკავშირებული პრობლემები: უშუალოდ ამ ამოცანით განსაზღვრული საპროექტო დავალების შესრულებასთან დაკავშირებით ჯგუფს არ ჰქონია არანაირი ხელისშემშლელი პრობლემა. თუმცა, „ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ლაბორატორიის“ არარსებობამ ხელი შეგვიშალა ავტომატურად გაფართოებადი მართლმწერის უფრო დახვეწილი ვერსიის აგებაშიც.

3. ხელშეკრულების მიხედვით ამოცანა N3-ით დაგეგმილი იყო ქართულ-მათემატიკური და ქართულ-ინგლისურ-გერმანული სემანტიკურად ორმხრივ მთარგმნელი სისტემის გაფართოებადი ანუ სწავლებადი ინტერნეტ ვერსიის შექმნა და პირობითად www.geotarnslator.com მისამართზე განთავსება. - ამ ამოცანით დაგეგმილი სამუშაოები

განხორციელდა სრულად. ამასთან, ამ ამოცანასთან დაკავშირებითაც, ჯგუფი გასცდა საპროექტო დავალებასა და გუგლის მთარგმნელი სისტემის ქართული გაფართოების სახით უკვე მოამზადა პირველი ქართული ხმიდან ხმაზე მთარგმნელი ექსპერიმენტული სისტემა, რომელიც ღია ფორმით განთავსებულია <http://geoanbani.com/> მისამართზე.

კაცი წერს ლამაზ წერილს.	A man is writing a beautiful letter.
-------------------------	--------------------------------------

From: Georgian
To: English
Translate

კაცი წერს ლამაზ წერილს.	Der Mann schreibt den schönen Brief.
-------------------------	--------------------------------------

From: Georgian
To: German
Translate

სურათი 4-5

მე-4-ე და მე-5-ე სურათებით წარმოდგენილია ქართულ-ინგლისური და ქართულ-გერმანული მთარგმნელი სისტემები მოქმედებაში. კერძოდ, როგორც ხედავთ, ჩვენმა მთარგმნელმა წინადადება „**კაცი წერს ლამაზ წერილს.**“ თარგმნა როგორც ინგლისურად („**A man is writing a beautiful letter.**“), ისე გერმანულად (**Der Mann schreibt einen schönen Brief.**“).

A man is writing a beautiful letter.	Der Mann schreibt den schönen Brief.
--------------------------------------	--------------------------------------

From: English
To: German
Translate

სურათი 6

Der Mann schreibt den schönen Brief.	კაცი წერს ლამაზ წერილს.
--------------------------------------	-------------------------

From:

To:

სურათი 7

მე-6-ე და მე-7-ე სურათებით წარმოდგენილია შებრუნებული ანუ ინგლისურ-ქართული და გერმანულ-ქართული მთარგმნელი სისტემები მოქმედებაში.

Der Mann schreibt den schönen Brief.	A man is writing a beautiful letter.
--------------------------------------	--------------------------------------

From:

To:

A man is writing a beautiful letter.	Der Mann schreibt den schönen Brief.
--------------------------------------	--------------------------------------

From:

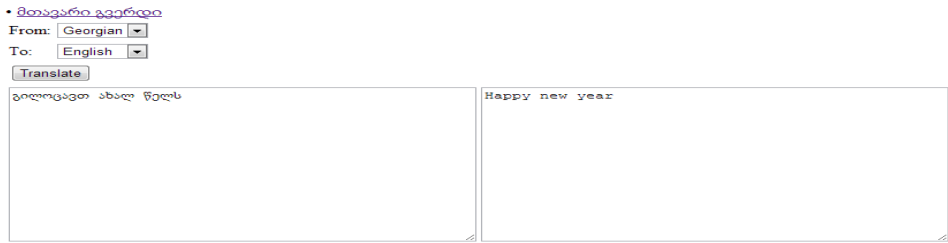
To:

სურათი 8-9

მე-8-ე და მე-9-ე სურათებით წარმოდგენილია გერმანულ-ინგლისური და ინგლისურ-გერმანული მთარგმნელი სისტემები, რაც იყო და არის ჩვენი კვლევის არაპირდაპირი, არამედ გვერდითი შედეგი, რადგან, რა თქმა უნდა, ჩვენი მიზანი არ არის ინგლისურ და გერმანულ ენებს შორის ისედაც მაღალი მთარგმნის ხარისხის აწევა.

ხაზგასასმელია, და უფრო მეტიც, საგანგაშოა, რომ, სამწუხაროდ, ამ ჯერ-ჯერობით ერთადერთი სემანტიკური ანუ წესებზე დამყარებული ქართული მთარგმნელი სისტემის მოქმედების ხარისხი და არეალი ჯერ კიდევ ძლიერ შეზღუდულია. თუმცა, ხაზგასასმელია ისიც, რომ გამოთვლითი ანუ მათემატიკური ლინგვისტიკის ფარგლებში ქართული ენა მაღალი ხარისხით მომუშავე მთარგმნელი სისტემის ასაგებად აუცილებელი თვალსაზრისებით სრულიად არასაკმარისად არის შესწავლილი, რაც ვითარების ზემოაღნიშნული საგანგაშობის ხარისხს კიდევ უფრო მაღალს ხდის.

ამასთან, როგორც უკვე აღვნიშნეთ, ამ ამოცანასთან დაკავშირებითაც, ჯგუფი გასცდა საპროექტო დავალებასა და გუგლის მთარგმნელი სისტემის ქართული გაფართოების სახით უკვე მოამზადა პირველი ქართული ხმიდან-ხმაზე მთარგმნელი ექსპერიმენტული სისტემა, რომელიც ღია ფორმით განთავსებულია <http://geoanbani.com/> მისამართზე. ეს ექსპერიმენტული სისტემა გუგლის მთარგმნელთან ერთად ეყრდნობა ამავე პროექტის N1 და N5 ამოცანების ფარგლებში ჩვენს მიერვე შემუშავებულ საინტერნეტო სისტემებს. ხაზგასასმელია ისიც, რომ ჯერ კიდევ, იმის გამო, არ არსებობს მაღალხარისხიანი ქართული მეტყველების ამომცნობი, ეს ხმიდან-ხმაზე მთარგმნელი ექსპერიმენტული სისტემა ქმედითუნარიანია მეტად შეზღუდული ლექსიკონის ფარგლებში.



© დამზადებულია საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების ცენტრში



© დამზადებულია საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების ცენტრში

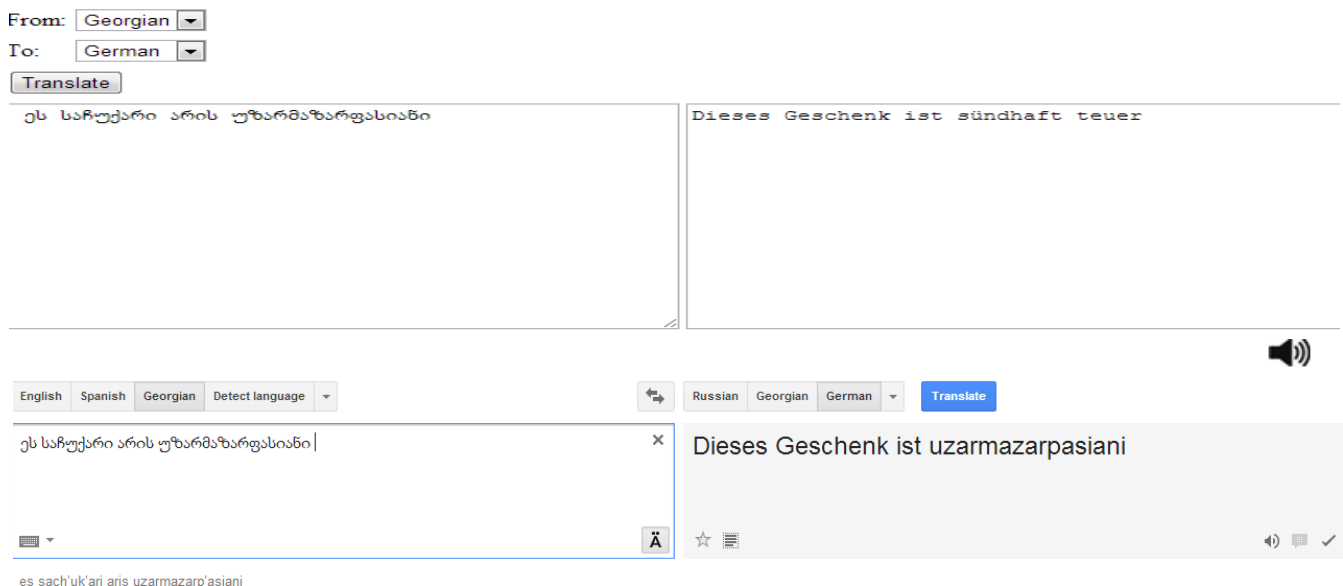
სურათი 10-11

10-11 სურათებზე გამოსახულია საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრში დამზადებული (ავტორები: კ.ფხაკაძე, გ.ჩიჩუა მ.ჩიქვინიძე) პირველი და ჯერ-ჯერობით ერთადერთი ქართული ხმიდან-ხმაზე მთარგმნელი ექსპერიმენტული სისტემა მოქმედებაში. პირველ (სურათი 10), შესაბამისად მეორე (სურათი 11), შემთხვევაში სისტემამ „გაისმინა“ რა მომხმარებლის მიერ ნათქვამი ფრაზა „გილოცავთ ახალ წელს“, შესაბამისად „გილოცავთ“,

მოგვაწოდა მისი როგორც წერილობითი, ისე ხმოვანი თარგმანი ინგლისურად, შესაბამისად რუსულად. გარდა ინგლისურისა და რუსულისა სისტემა ქართულიდან თარგმნის აგრეთვე გერმანულ, ფრანგულ, ესპანურ, იტალურ და თურქულ ენებშიც.

ხაზგასასმელია, და უფრო მეტიც, საგანგაშოა, რომ, სამწუხაროდ, ჯერ-ჯერობით ამ ერთადერთი ქართული ხმიდან-ხმაზე მთარგმნელი ექსპერიმენტული სისტემის მოქმედების ხარისხი და არეალი ჯერ კიდევ ძლიერ შეზღუდულია. თუმცა, ხაზგასასმელია ისიც, რომ გამოთვლითი ანუ მათემატიკური ლინგვისტიკის ფარგლებში ქართული ენა ხმიდან-ხმაზე მაღალი ხარისხით მომუშავე მთარგმნელი სისტემის ასაგებად აუცილებელი თვალსაზრისებით სრულიად არასაკმარისად არის შესწავლილი, რაც ვითარების ზემოაღნიშნული საგანგაშობის ხარისხს კიდევ უფრო მაღალს ხდის.

გარდა ამისა, ქვემოთ, 12-17 სურათებზე გამოსახულია საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრში დამზადებული (ავტორები: კ.ფხაკაძე, მ.ჩიქვინიძე, გ.ჩიჩუა, ი.ბერიაშვილი) გუგლის ქართულ-გერმანული მთარგმნელის კიდევ ერთი ექსპერიმენტული გაფართოება, რითაც ჯგუფი ასევე გასცდა ამოცანა N3-ით დასახულ საპროექტო მიზნებს. კერძოდ, სურათ 12-ზე ნაჩვენებია ჩვენი თარგმანი წინადადებებისა „ეს საჩუქარი არის უზარმაზარფასიანი“, ხოლო სურათ 13-ზე ნაჩვენებია იგივე წინადადების თარგმანი გუგლის მთარგმნელით. - როგორც ხედავთ გუგლის მთარგმნელში „უზარმაზარფასიანი“ არ ითარგმნება, გუგლის მთარგმნელის ჩვენმა გაფართოებამ კი „უზარმაზარფასიანი“ თარგმნა.



სურათი 12-13

ქვემოთ, მე-14-ე სურათზე, ნაჩვენებია 20 ქართული სიტყვა, რომელსაც გუგლის მთარგმნელი გერმანულად ვერ თარგმნის, მე-15-ე სურათზე კი ნაჩვენებია ჩვენი გაფართოების მიხედვით ამ ოცი სიტყვის გერმანული თარგმანი.

Georgian English Spanish Detect language

German English Russian Translate

დაჭყანული ჩავჩიჩინებდი შეამხანაგებას გაქნლი მიმიწევს დაშამულ იყოჩაღებენ მადგება ცმუკვას ულევი მიშლიანო შეკვილევა ფართხაფურთხის გამიელვა შემომმახეს შემოსროლილი დასუნამოებული გააცილიდა თავგადახოტრილს თავყანისმცემნი.

Dachqanuli Chavchichinebdi Sheamkhanagebas Gaknili Mimitsevs Dashkhamul Iqochagheben Madgeba Tsmukvas Ulevi Mishliano Shekivleba Partkhapurtkhis Gamielva Shemomdzakhes Shemosroliili Dasunamoebuli Gaatslida Tavgadakhotrils Taqvanismtsemni.

From: Georgian To: German Translate

დაჭყანული ჩავჩიჩინებდი შეამხანაგებას გაქნლი მიმიწევს დაშამულ იყოჩაღებენ მადგება ცმუკვას ულევი მიშლიანო შეკვილევა ფართხაფურთხის გამიელვა შემომმახეს შემოსროლილი დასუნამოებული გააცილიდა თავგადახოტრილს თავყანისმცემნი.

grimassiert zureden die Partnerschaft Heuchler wird schieben giftig werden eifrig nützt die Hektik ohne Ende verbieten mir kreischen der Hektik ist in mir aufgeblitzt haben mich aufgerufen reingeworfen parfümiert entfernte Kahlgescherten Fans.

სურათი 14-15

კიდევ უფრო გასაგები რომ იყოს რაზეა საუბარი, შეადარეთ ერთმანეთს მე-16-ე და მე-17-ე სურათებით მოცემული თარგმანები. ცხადია, რომ, ლექსიკურად, მე-17-ე სურათით მოცემული ჩვენი ქართულ-გერმანული თარგმანი უფრო სრულია, ვიდრე მე-16-ე სურათით მოცემული გუგლის თარგმანი.

Georgian English Spanish Detect language

German English Russian Translate

დაჭყანული ჩავჩიჩინებდი და ვთხოვდი შეამხანაგებას გაქნლი კაცია კარადას ის მიმიწევს დაშამულ სოკოს ჭამდა ცმუკვას ვერ ვიტან არ იყოჩაღებენ ულევი არის არაფერი შეკვილევა მე გავიგონე მეგობრებმა შემომმახეს თავყანისმცემნი იყვნენ ბევრნი.

Dachqanuli chavchichinebdi und fragen sheamkhanagebas Gaknili Mann. Er mimitsevs Schrank. Dashkhamul Pilz Ate. Ich hasse tsmukvas. Nicht iqochagheben. Ulevi Ist nichts. Shekivleba ich gehört habe. Freunde Shemomdzakhes. Taqvanismtsemni waren viele.

From: Georgian To: German Translate

დაჭყანული ჩავჩიჩინებდი და ვთხოვდი შეამხანაგებას გაქნლი კაცია კარადას ის მიმიწევს დაშამულ სოკოს ჭამდა ცმუკვას ვერ ვიტან არ იყოჩაღებენ ულევი არის არაფერი შეკვილევა მე გავიგონე მეგობრებმა შემომმახეს თავყანისმცემნი იყვნენ ბევრნი.

grimassiert zureden und fragen die Partnerschaft Heuchler Mann. Er wird schieben Schrank giftig ab den Pilz. Ich hasse die Hektik. Nicht werden eifrig ohne Ende ist nichts kreischen ich gehört habe haben mich aufgerufen Freunde Fans waren viele.

სურათი 16-17

თუმცა, ცხადია ისიც, რომ ამ ორივე თარგმანში მრავალი სისუსტეა, რაც არის არა ჩვენი გაფართოების მიზეზით, არამედ თავად გუგლის მთარგმნელის სუსტი მხარეების გამო. აქ საუბარია გუგლის სტატისტიკურ მიდგომებზე დამყარებული მთარგმნელის იმ

სუსტ მხარეებზე, რომელთაც კარგ დემონსტრირებას უკეთებს ქვემოთ წარმოდგენილი სურათები:



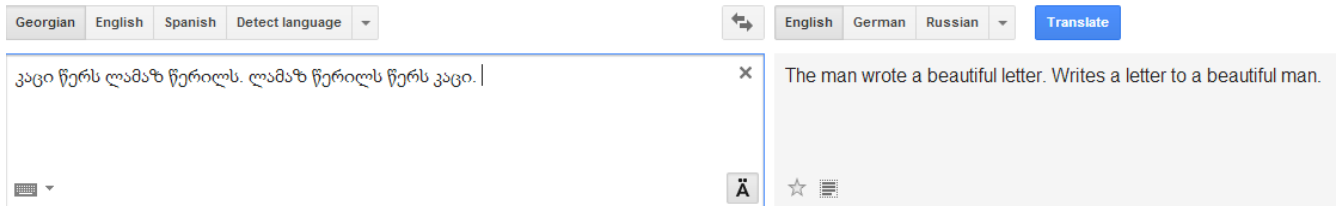
From: Georgian
To: English
Translate

სურათი 18

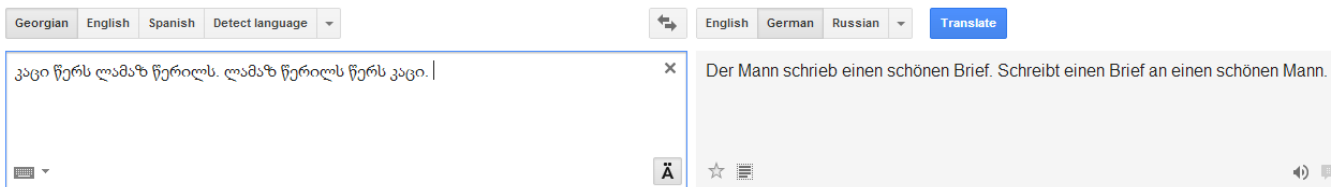


From: Georgian
To: German
Translate

სურათი 19



სურათი 20



სურათი 21

მე-18-ე და მე-19-ე სურათებზე ნაჩვენებია როგორ მუშაობს ჩვენი სემანტიკური ანუ წესებზე დამყარებული ანუ მათემატიკური ენის საშუაშავლო მონაწილეობით მოქმედი ქართულ-ინგლისურ-გერმანული მთარგმნელი ორ წინადადებიან ტექსტზე. კერძოდ,

როგორც ხედავთ, ჩვენმა მთარგმნელმა ტექსტი „*კაცი წერს ლამაზ წერილს. ლამაზ წერილს წერს კაცი.*“ ინგლისურად და გერმანულად ორ ერთიდაიგივე წინადადებით აგებულ ტექსტად თარგმნა, რაც, ცხადია, ასეც უნდა ყოფილიყო. 20-ე და 21-ე სურათებზე ნაჩვენებია ნაჩვენებია იგივე ორწინადადებიანი ქართული ტექსტის გუგლისეული ინგლისური („*A man writes a beautiful letter. Writes a letter to a beautiful man.*“) და გერმანული (*Der Mann schrieb einen schönen Brief. Schreibt einen Brief an einen schönen Mann.*) თარგმანები, რაც, რა თქმა უნდა, ვერანაირ კრიტიკას ვერ უძლებს.

ამოცანა N3-ით განსაზღვრული საპროექტო დავალების შესრულებასთან დაკავშირებული პრობლემები: უშუალოდ ამ ამოცანით განსაზღვრული საპროექტო დავალების შესრულებასთან დაკავშირებით ჯგუფს არ ჰქონია არანაირი ხელისშემშლელი პრობლემა. თუმცა, „ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ლაბორატორიის“ არარსებობამ ხელი შეგვიშალა როგორც ქართული მკითხველის (ანუ ქართული კომპიუტერული ყელის), ისე ქართული მეტყველების ამომცნობის (ანუ ქართული კომპიუტერული ყურის) უფრო დახვეწილი ვერსიების აგებაში, რამაც, მთლიანობაში, ხელი შეგვიშალა აგრეთვე ამ ამოცანით მიღწეული შედეგის შემდგომ გაუმჯობესებაშიც.

4. ხელშეკრულების მიხედვით ამოცანა N4-ით დაგეგმილი იყო ქართული ლოგიკური ამოცანების ამომხსნელი და შემმოწმებელი ექსპერიმენტული ექსელ სისტემის გაფართოებადი ანუ სწავლებადი ინტერნეტ ვერსიის შექმნა და პირობითად www.georeasoner.com მისამართზე განთავსება. - ამ ამოცანით დაგეგმილი ქართული ლოგიკური ამოცანების ამომხსნელი და შემმოწმებელი ექსპერიმენტული სისტემის ინტერნეტ ვერსია უახლოეს მომავალში განთავსდება <http://geoanbani.com/> მისამართზე.

ამოცანა 4-ით განსაზღვრული საპროექტო დავალების შესრულებასთან დაკავშირებული პრობლემები: ამ ამოცანით განსაზღვრული საპროექტო დავალების შესრულებასთან დაკავშირებით ჯგუფს არ ჰქონია არანაირი ხელისშემშლელი პრობლემა. თუმცა, „ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ლაბორატორიის“ არარსებობამ ამ შემთხვევაშიც ხელი შეგვიშალა შედეგების უფრო ჩქარი ნაბიჯებით წინ წაწევაში

5. ხელშეკრულების მიხედვით ამოცანა N5-ით დაგეგმილია შეზღუდულ სიტყვარიანი ქართული მეტყველების ამომცნობი სისტემის გაფართოებადი ანუ სწავლებადი ინტერნეტ ვერსიის შექმნა და პირობითად www.geospeechlistener.com მისამართზე განთავსება. - ამ ამოცანით დაგეგმილი სამუშაოები განხორციელდა სრულად. ამასთან, ამ შემთხვევაშიც ჯგუფი გასცდა საპროექტო დავალებას და უკვე შეიქმნა თვითსწავლისა და სწავლების პრინციპით აგებული შეზღუდულ სიტყვარიანი ქართული მეტყველების ამომცნობი სისტემის გაფართოებადი ანუ სწავლებადი პირველი და ჯერ-ჯერობით ერთადერთი ინტერნეტ ვერსია, რომელიც ღია სახით იძებნება <http://geoanbani.com/> მისამართზე.

დააჭირეთ Allow ღილაკს და თქვით ფრაზა



სურათი 22

22-ე სურათზე გამოსახულია საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ცენტრში დამზადებული (ავტორები: კ.ფხაკაძე, გ.ჩიჩუა) პირველი და ჯერ-ჯერობით ერთადერთი თვითსწავლისა და სწავლების პრინციპით აგებული შეზღუდულ სიტყვიერი ქართული მეტყველების ამომცნობი სისტემა.

ხაზგასასმელია, და უფრო მეტიც, საგანგაშოა, რომ, სამწუხაროდ, ამ ჯერ-ჯერობით ერთადერთი ქართული მეტყველების ამომცნობი სისტემების მოქმედების ხარისხი და არეალი ჯერ კიდევ ძლიერ შეზღუდულია. თუმცა, თავიც რომ დავანებოთ ზემოხსენებულ ლაბორატორიას, რომლის ფორმირების შემდეგ ჩვენ აქ შედეგების შემდგომი გაუმჯობესების ბევრად უფრო მაღალი შანსები გვეძლევა, ხაზგასასმელია ისიც, რომ გამოთვლითი ანუ მათემატიკური ლინგვისტიკის ფარგლებში ქართული მეტყველება ამომცნობი სისტემების ასაგებად აუცილებელი თვალსაზრისებით სრულიად არასაკმარისად არის შესწავლილი, რაც ვითარების ზემოაღნიშნული საგანგაშობის ხარისხს კიდევ უფრო მაღალს ხდის.

ამოცანა N5-ით განსაზღვრული საპროექტო დავალების შესრულებასთან დაკავშირებული პრობლემები: უშუალოდ ამ ამოცანით განსაზღვრული საპროექტო დავალების შესრულებასთან დაკავშირებით ჯგუფს არ ჰქონია არანაირი ხელისშემშლელი პრობლემა. თუმცა, „ქართული ენის ტექნოლოგიების სასწავლო-სამეცნიერო ლაბორატორიის“ არარსებობამ, რის გარეშეც, როგორც უკვე აღინიშნა, ჩვენ ვერ ვახერხებთ უფრო ხარისხიანი ქართული მკითხველის ანუ ქართული კომპიუტერული ყელის აგებას, ხელი შეგვიშალა აგრეთვე ქართული მეტყველების ამომცნობის ანუ ქართული კომპიუტერული ყურის უფრო ხარისხიანი ვერსიების აგებაშიც.

6. ხელშეკრულების მიხედვით ამოცანა N6-ით, რომელიც ჯგუფმა დამატებით საპროექტო მიზნად აიღო, დაგეგმილი იყო ქართული სამწერლებო და სამეტყველო ენების თვითგაფართოებადი ანუ თვითსწავლებადი ინტელექტუალური (მოაზროვნე) კორპუსის, შემდგომში მოკლედ - ქართული ინტელექტუალური კორპუსის საცდელი ვერსიის შექმნა. - ამ ამოცანით დაგეგმილი სამუშაოები განხორციელდა სრულად. ამასთან, ამ შემთხვევაშიც

ჯგუფი გასცდა საპროექტო დავალებას და უკვე ააგო ქართული ინტელექტუალური კორპუსის საცდელი ვერსია მასში ჩადგმული საკმაოდ მძლავრი ინტელექტუალური პროცედურებითურთ.

კორპუსის ძირითადი გვერდი, რომელიც ღია წვდომის ქვეშაა, განთავსებულია <http://geoanbani.com/Corpus/> მისამართზე. ამ მისამართის გააქტიურებთ ევრანზე გამოსული გვერდის მარჯვენა კუთხეში განთავსებულია მთვლელი, რომლის პირველი ხაზში არსებული ჩანაწერი გვიჩვენებს მიმდინარე მომენტში კორპუსში შემოსულ სიტყვათა ჯამურ რაოდენობას, მეორე ხაზის ჩანაწერი გვიჩვენებს კორპუსში უკვე არსებულ განსხვავებულ სიტყვათა რაოდენობას, მესამე ხაზის ჩანაწერი კი გვიჩვენებს წინა 24 საათში შემოსული ახალი სიტყვების რაოდენობას. ამგვარად, სურათი 23-დან ჩანს, რომ 24 საათის განმავლობაში, ანუ 2014 წლის 17 აპრილის 15:30 - საათიდან 2014 წლის 18 აპრილის 15:30 საათამდე:

1. კორპუსში შემოსულ სიტყვათა ჯამური რაოდენობა გაიზარდა 14486 სიტყვით. მართლაც, $124278770 + 14486 = 124293256$;
2. კორპუსში შემოსული ახალი სიტყვების რაოდენობა გაიზარდა 124 სიტყვით. მართლაც, $2153023 + 124 = 2153147$.

2014 წლის 17 აპრილის 15:30 საათისათვის

სულ შემოსული:	124293256
განსხვავებული:	2153147
ბოლო დღეს:	124

2014 წლის 16 აპრილის 15:30 საათისათვის

სულ შემოსული:	124278770
განსხვავებული:	2153023
ბოლო დღეს:	132

სურათი 23

სურათი 24-ზე ასახულია ინფორმაცია რასაც კორპუსი იძლევა სიტყვაზე „პატრიოტი“. სურათიდან ირკვევა, რომ ამ სიტყვის ფარდობითი სიხშირეა 0.0008%, სიხშირე - 950. ასევე, კორპუსი გვაძლევს ამ სიტყვის შემცველ წინადადებებს მათზე მიდგმული ლინკებით, რომელთა გააქტიურებით გვეძლევა ამ წინადადებების შემცველი ტექსტი. წინადადებების გარდა, კორპუსი იძლევა დასამუშავებლად მიცემული სიტყვის როგორც მარჯვენა, ისე მარცხენა, ასევე ორმხრივ კონტექსტებს.

სიტყვა: პატრიოტი | ინფორმაციის მოძიება

ქართული: პატრიოტი – ფარდობითი სიხშირე: 0.0008% რაოდენობა: 950

ინგლისური: Patriots | სტატუსი: PERFECT

გერმანული: Patriots | სტატუსი: PERFECT

წინადადებები:
 ყოველი წამდელი პატრიოტი კოსმოპოლიტია ისე, როგორც ყოველი გონიერი კოსმოპოლიტი (და არა ჩვენებური) პატრიოტია. [ჩხელი](#)
 პოეტის თითქმის მათი შრომების მომწიფლად და უშუალოდ გამოცდის მათი გამარჯვებით გამოწვეულ აღფრთოვანებას, თითქმის ყრუბა საუკუნეთა მანძილი და მზუნხარე პატრიოტი
 თამარ მეფის საბეჭო დარბაზის სილამაზეს ისევე გურჩეულად და ახლოს განიცდის, როგორც გაზაფხულის სიდიადეს. [ჩხელი](#)
 ჩვენ კი, მის მეგობრებს, რომელთაც პოეტის მზუნხარე სტრუქტურებთან ერთად მისი ფართო, ცისფერი თვალის ამაყი გამოხედვაც გვახსოვს, მარად დაგვირგება პატრიოტი პოეტის
 ნათელი ხსოვნა. [ჩხელი](#)

კონტექსტები:
 მზუნხარე პატრიოტი
 დაგვირგება პატრიოტი
 ბევრი პატრიოტი
 დიდი პატრიოტი

პატრიოტი თამარ
 პატრიოტი პოეტის

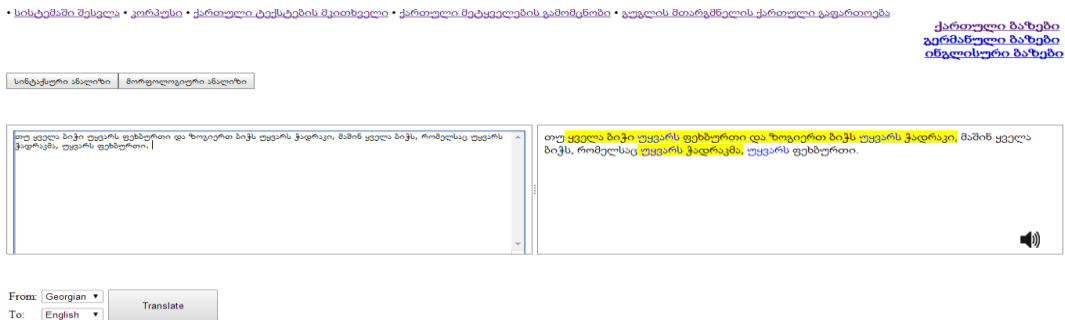
მზუნხარე პატრიოტი თამარ
 დაგვირგება პატრიოტი პოეტის

© გამოყენებულია საქართველოს ტექნიკური უნივერსიტეტის ქართული ენის ტექნოლოგიების ცენტრში

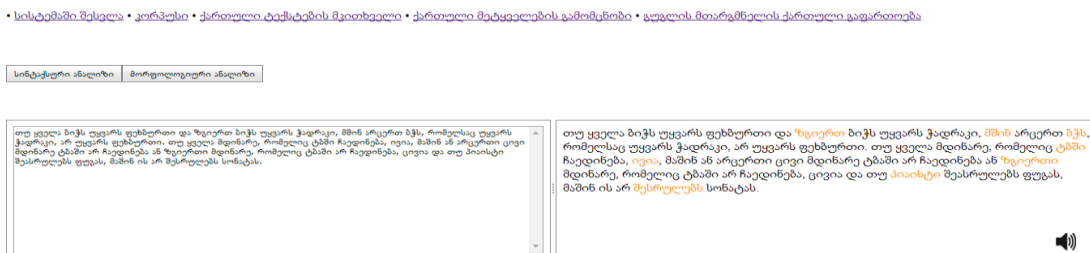
სურათი 24

გარდა ამისა, კორპუსი გუგლის მთარგმნელის მიხედვით გვაძლევს სიტყვის როგორც ინგლისურ, ისე გერმანულ თარგმანს როგორც წერითი, ისე ხმოვანი ფორმით (იხ. სურათი 24). ასევე, ქართული ენის ხაზში არსებული დინამიკის იკონის გააქტიურებით მომხმარებელს კორპუსში ჩადგმული ჩვენს მიერვე შემუშავებული მკითხველი სისტემის მეშვეობით შეუძლია მოისმინოს ქართული სიტყვის წაკითხული ანუ სინთეზირებული ვარიანტი. ანუ, ფაქტიურად, ვეყრდნობით რა გუგლის მთარგმნელ სისტემას, კორპუსში ჩადგმული სახით უკვე აგებულია ქართულ-ინგლისურ-გერმანული ხმოვანი ლექსიკონი.

ამგვარად, ხაზგასასმელია, რომ კორპუსი ამ მასში ჩადგმული მკითხველი სისტემის მეშვეობით სიტყვების წერით ფორმებთან ერთად ანუ მათთან დაწყვილებული სახით ინახავს მათ ხმოვან ანუ სინთეზირებულ ფორმებსაც. ეს ჩვენ, ერთი მხრივ, გვჭირდება სტუ ქართული ენის ტექნოლოგიების ცენტრში უკვე შემუშავებული ქართული მეტყველების გამომცნობი სისტემის მუშაობის ხარისხის შემდგომი არსებითი ხასიათის გაუმჯობესებისათვის, მეორე მხრივ კი, ეს ჩვენ უკვე გამოვიყენოთ. ანუ, იმაზე დაყრდნობით, რომ კორპუსში სიტყვების წერით ფორმებთან უკვე დაწყვილებულია მათი სინთეზირებული ფორმებიც, ჩვენ უკვე ავაგეთ ქართული ვებ-გვერდების მკითხველი პირველი და ჯერ-ჯერობით ერთადერთი სისტემა, რომელზედაც ზემოთ უკვე გვქონდა საუბარი.



სურათი 25



სურათი 26

ზემოთ 25-ე სურათზე თქვენ ხედავთ ჩვენი საიტის შიგა ანუ არაღიაწვდომის პირველ გვერდს, 26-ე სურათზე კი - ღიაწვდომის პირველი გვერდია წარმოდგენილი, რომელზეც დახურული წვდომის არაყველა ფუქნცუიონირებადი ლინკია გამოტანილი. ამასთან, 25-ე სურათზე ნაჩვენებია როგორ მოქმედებს კ.ფხაკაძის ქართული ენის ლოგიკური გრამატიკის საფუძვლებზე დაყრდნობით შემუშავებული ავტომატურად

გაფართოებადი ქართული სინტაქსური მართლმწერი ანუ მანალიზებული სისტემა, რომელიც, როგორც უკვე აღვნიშნეთ, ჯერ-ჯერობით, ერთად-ერთი ამ ტიპის სისტემა ქართულისათვის, 26-ე ნახაზზე კი წარმოდგენილია ჩვენ კორპუსზე დამყარებული ქართული თვითსწავლებადი - ანუ თვითგაფართოებადი ორთოგრაფის მოქმედების შედეგი. ანუ, ორთოგრაფმა, გამომდინარე იმ ცოდნიდან, რაც მას ტექსტის გასწორების დროისთვის ჰქონდა, გაანალიზა შემოსული ტექსტი და წითლად შეაფერადა მისი ვარაუდით არასწორი სიტყვები.

განსაკუთრებით ხაზგასასმელია, რომ კორპუსში უკვე ჩადგმულია ქართულ ენაში არსებითი და ზედსართვი სახელებისა და ზმნების გამომცნობი ინტელექტუალური პროცედურები, რომლებიც უკვე მოქმედებენ საკმაოდ ფართოდ და საკმაოდ მაღალი ალაბათური სიზუსტით. თანაც, ეს პროცედურები ამ უკვე გამოცნობილი ფორმებით სინტაქსური ანალიზატორის ბაზებზე მისაერთებლად ვარგისი ფორმით ანუ უკვე სინტაქსურად აღწერილი ფორმით ავტომატურად აგენერირებენ გამომცნობილი სახელების სხვადასხვა ბრუნვით ფორმებსა და, ასევე, გამომცნობილი ზმნების სხვადასხვა უღლებით ფორმებს, რითაც ჩვენი მართლმწერ-ანალიზატორი სისტემა, როგორც უკვე აღინიშნა, კორპუსზე დაყრდნობით ავტომატურად ფართოვდება ანუ ავტომატურად სწავლობს ანუ ავტომატურად იფართოვებს მასში უკვე მანანდე არსებულ ქართულ ენობრივი ცოდნას.

დაბოლოს, ხაზგასასმელია ისიც, რომ უკვე დაწყებულია მუშაობა ჩვენს მიერვე შემუშავებული (კ.ფხაკაძე, მ.ჩიქვინიძე, ი.ბერიაშვილი, ა.მასხარაშვილი) წესებზე დამყარებული ქართულ-გერმანულ-ინგლისური მთარგმნელი სისტემების კორპუსზე მიერთების მიმართულებით. კერძოდ, უკვე მუშავდება ამ მთარგმნელი სისტემების ცოდნის ავტომატურად გამაფართოებელი ინსტრუმენტების ამგები ხერხები და მეთოდები.

გრანტის სახსრებით შეძენილი მცირეფასიანი საგნებისა და ძირითადი საშუალებების ნუსხა:

1. 2014 წლის 14 აპრილს სტუ გ.ნიკოლაძის სხდომათა დარბაზში ჩატარდა ქართული ენის დღისადმი მიძღვნილი ერთდღიანი კონფერენცია-სემინარი „ქართული ენის ტექნოლოგიური ანბანი - 21-ე საუკუნის ერთ-ერთი უმთავრესი გამოწვევა“. ამასთან, კონფერენციის მომზადების მიზნით დაიბეჭდა გასაკრავი და დასარიგებელი საინფორმაციო ბუკლეტები (10+250=260) სულ 260 ცალი.

სტუ № 048-13 საგრანტო პროექტის ხელმძღვანელი,
სტუ ქართული ენის ტექნოლოგიების ცენტრის
დირექტორი, პროფ. /კონსტანტინე ფხაკაძე/