

**კონცეფცია ეტიმოლოგიური კიბაბის კომპიუტერული  
სისტემისთვის**

გელა ლვინუაძე  
საქართველოს ტექნიკური უნივერსიტეტი

**რეზიუმე**

თანამედროვე მეთოდებზე დაფუძნებულმა კვლევებმა მოგვცა ასეთი შედეგი - თავდაპირველი სიტყვების უმრავლესობისათვის ამოსავალი ცნებაა წყალი. ამ მიგნებაზე დაყრდნობით, დადგინდა ეტიმოლოგიური თვალსაზრისით რიგი „უიმედო“ ქართული სიტყვების წარმოქმნის გზები. სიტყვების ეტიმოლოგიის და რიგი სხვა ძიებების ავტომატიზების მიზნით, შემუშავდა კომპიუტერული ექსპერტული სისტემის კონცეფცია.

**საკვანძო სიტყვები:** კომპიუტერული ექსპერტული სისტემები. ეტიმოლოგიების ძიება.

**1. შესავალი**

უკანასკნელი წლებში მთელს მსოფლიოში შეიმჩნევა ცხოველი ინტერესი ენების საყაროსადმი, რაზეც მეტყველებს თუნდაც მხოლოდ ერთი შემდეგი ფაქტი:

დამატებით იმისა, რომ ვიკიპედიაში მსოფლიოში არსებული მეტნაკლებად მნიშვნელოვანი თითქმის ყველა ენისადმი არის მიძღვნილი სტატია, ბოლო ხანებში გაჩნდა მისი განშტოება - ვიკილექსიკონის სახელით ცნობილი საიტების მთელი წყება მსოფლიოს განვითარებული ქვეყნების ენებზე [1-4]. თითოეული მათგანი წარმოადგენს წარმოადგენს მულტიფუნქციურ მრავალენოვან ლექსიკონსა და თეზაურუსს (სამუშაო ენის დასახელება შემოკლებული სახით ფიგურირებს საიტის მისამართში). ისევე, როგორც ვიკიპედიის სხვა პროექტების, ასევე ვიკილექსიკონის შევსებასა და მისი ღირსება-ნაკლოვანებების განსჯა-შეფასებაში მონაწილეობის მიღება შეუძლია ნებისმიერ მსურველს მსოფლიოს ყველა კუთხიდან. აღნიშნული სახის საიტებმა სწრაფად მოიპოვეს პოპულარობა, მით უფრო, როდესაც ორიგინალის [1] შესაძლებლობების მისადაგება ადვილად შეიძლება მოხდეს, ფაქტობრივად, ნებისმიერი კონკრეტული ენისათვის. მაგალითად, ვიკილექსიკონის რუსული ენის ვარიანტი, რომელიც ფუნქციონირებს 2004 წლის დასაწყისიდან, დღეისათვის უკვე მოიცავს 300 ათასამდე სტატიას სიტყვების, სიტყვათწარმოქმნელი ერთეულების და სიტყვათშეთანხმებების შესახებ მსოფლიოს 300-ზე მეტ ენაზე [2], ხოლო თვით ინგლისურისათვის კი სტატიების რაოდენობა დღეისათვის 2,5 მლნ-ს აღწევს, ენების რიცხვი კი - 400-ს აჭარბებს [1]. აქვე აღვნიშნავთ, რომ, სამწუხაროდ, საქართველოს ჯერჯერობით ამ მიმართულებით ნაბიჯები არ გადაუდგამს.

ვიკილექსიკონ-საიტების მეშვეობით შესაძლებელია ჩვენთვის საინტერესო სიტყვისათვის პარალელები მოვიძიოთ, როგორც მოცემულ და მის მონათესავე ენებზე, ასევე - რიგ შემთხვევებში ენების იერარქიულ ხეზე სამუშაო ენისაგან დიდად დაცილებულ შტოებზეც.

გარდა ზემოთ აღნიშნულისა, კომპიუტერული პროგრამა გვთავაზობს რიგი სხვა დახმარებების გაწევასაც. კერძოდ, არჩეული სიტყვისათვის შესაძლებელი გახლავთ ენის თეზაურუსიდან მივიღოთ შემდეგი სახის ინფორმაცია [5]:

1. ენის დასახელება
  - 1.1 მორფოლოგიური და სინტაქსური თვისებები
  - 1.2 გამოთქმა
  - 1.3 სემანტიკური თვისებები
    - 1.3.1 მნიშვნელობა
    - 1.3.2 სინონიმები
    - 1.3.3 ანტონიმები
    - 1.3.4 ჰიპერონიმები
    - 1.3.5 ჰიპონიმები
  - 1.4 მონათესავე სიტყვები
  - 1.5 ეტიმოლოგია
  - 1.6 ფრაზოლოგიური ზმები და დამკვიდრებული სიტყვათშეხამებები

ვუჩვენოთ ვიკილექსიკონით სარგებლობის მაგალითი. დავუშვათ, გვინტერესებს ინფორმაცია მივიღოთ რუსული სიტვა -ს შესახებ - მოვიძიოთ მისთვის სინონიმები, ანტონიმები, კავშირები მონათესავე თუ არამონათესავე ენებზე არსებულ სიტყვებთან და სხვ.

ბროუზერის სამისამართო სტრიქონში ვკრეფთ: <http://ru.wiktionary.org/wiki/>\_\_\_\_\_. ეკრანზე აისახება საჭირო ინფორმაცია. ავირჩევთ შესაბამის პუნქტს და გავეცნობით მის შემცველობას.

უდავო ღირსებებთან ერთად ამ სისტემას გააჩნია მნიშვნელოვანი ნაკლი. კერძოდ, მასში არ გამოიყენება ის შესაძლებლობები, რომლის რეალიზებაც ძალუძთ კომპიუტერულ ექსპერტულ სისტემებს,

კერძოდ, ექსპერტული შეფასებების ავტომატიზება და პროგრამული პროცედურების ღონეზე დაყვანილი ხელოვნური ინტელექტის მეთოდების მეშვეობით ადამიანი-მანქანა დიალოგში გადაწყვეტილებების მიღება.

სწორედ აღნიშნული მიდგომისათვის შესაბამისი კონცეფციის შემუშავება შეადგენს წამოღვენილი ნაშრომის მიზანს.

## 2. ძირითადი ნაწილი

სანამ ლიგვისტიკური ამოცანების გადასაწყვეტად და კერძოდ, ეტიმოლოგიების ძიების პროცესისათვის განკუთვნილი კომპიუტერული სისტემის აგების კონცეფციის ჩვენ მიერ შემუშავებული კონცეფციის აღწერაზე გადავიდოდეთ, მანამდე მისთვის მყარი ნიადაგის მოსამზადებლად მიგვაჩნია, რომ საჭიროა განვიხილოთ ზოგიერთი ფუნდამენტური სახის თეორიული საკითხი და წარმოვადგინოთ ამ მიმართულებით საკუთარი ხედვაც.

არსებობს მეცნიერული ვარაუდი, რომ მსოფლიოში არსებული ენები ერთი, ძირეული ენისგან იღებენ სათავეს. აქვე აღვნიშნავთ, რომ ეს ვარაუდი სავსებით ეთანხმება ბიბლიურ შეხედულებებს. იყო მცდელობები, განესაზღვრათ ამ ჰიპოთეზური ენის ლექსიკა. ცხადია, ათასწლეულების მანძილზე ფრიად სახეცვლილი სიტყვებისათვის სრულებითაც არ არის ადვილი საქმე, მოძიებული იქნეს მათი საერთო ფესვები, მაგრამ გაცილებით რთულდება „პირველმისწინ“ დადგენა საწყისი ლექსიკის მეტ-ნაკლებად სავარაუდო ერთეულებისათვის.

წარუმატებლობებმა ამ გზაზე მეცნიერთა ნაწილს აფიქრებინებდა, რომ თავდაპირველი სიტყვებისთვის ეტიმოლოგიების ძიება, საერთოდ, უიმედო საქმეა. მათი აზრით, ეს ერთეულები, უბრალოდ, უმოტივაციო ცნებებია.

თუმცა ბოლო ხანებში ჩვენ მიერ გამოითქვა საპირისპირო აზრიც [6-7] – ამგვარი სიტყვების ეტიმოლოგიის დადგენა არცთუ ისე უპერსპექტივო საქმეა და სისტემური კვლევის მეთოდებზე დაყრდნობით, გამოკვეთილი იქნა რიგი ისეთი ცნებებისა, რომელთაც, სავარაუდოდ, დასაბამი მისცეს დღეს მსოფლიოში არსებულ ენათა ფრიად განტოტვილ ოჯახს.

კვლევების მომდევნო ეტაპზე მიზნად დავისახეთ, ცნებათა აღნიშნული წრიდან გამოგვეჩინა ის ტერმინი, რომლის მეშვეობითაც ამა თუ იმ ობიექტის, ქმედების თუ მოვლენის სახელდება, უპირველეს ყოვლისა, დასჭირდებოდა ადამიანს.

საზს ვუსვამთ შემდეგ გარემოებას - საუბარია არა რომელიმე კონკრეტული სიტყვის (რაიმე ბგერათკომპლექსის) მოძიებაზე, არამედ იმ პირველ ცნებაზე, რომელმაც სათავე დაუდო მეტყველებას.

დასმულ კითხვაზე პასუხის მისაღებად გამოვიყენეთ ექსპერტული შეფასებების, კერძოდ კი, დელფოსის მეთოდი. იგი, ინარჩუნებს რა ჯგუფური გამოკითხვის მეთოდებისთვის დამახასიათებელ ყველა ღირსებას, ხასიათდება რიგი სხვა დადებითი მხარეებითაც (ანონიმურობა, წინა ტურებში მიღებული შედეგები მაქსიმალური ეფექტიანობით გამოყენება, პასუხების სიმრავლისთვის სტატისტიკური მახასიათებლების - მედიანის და კვარტილების – განსაზღვრა).

კვლევები იმგვარ სფეროებში, როგორც ლინგვისტიკა გახლავთ, შეიძლება შევადართოთ გამოძიების პროცესს - აქაც ამა თუ იმ ვერსიის გასამარგებლად თუ უარსაყოფად ფაქტებსა და მტკიცებულებებს ხშირად სრულიად განსხვავებული წყაროებიდან მოიხმობენ, შემდეგ კი ხდება მათი გაანალიზება-შეჯერება და საბოლოო დასკვნის გამოტანა.

სწორედ ეს მიდგომა შევირჩიეთ “პირველი სიტყვის” მოძიებისთვისაც. წყაროებად გამოვიყენეთ: ბიბლია, უძველესი ხალხების მითოლოგიები, სხვადასხვა მეცნიერულ სფეროებში (ფიზიკა, ქიმია, ბიოლოგია) მიღებული შედეგები, ექსპერტთა გამოკითხვები.

დელფოსის მეთოდით ჩატარებული კვლევების უკვე მეორე ტურის შემდეგ კიდევ ერთხელ დავრწმუნდით აღნიშნული მეთოდისათვის დამახასიათებელ ეფექტიანობაში – პასუხები გადაადგილდა შემდეგი მტკიცებისაკენ:

ადამიანის მიერ სახელდებული პირველი ცნება არის წყალი.

ჰიპოთეზის ქმედითუნარიანობა გადავწყვიტეთ შეგვემოწმებინა ეტიმოლოგიური თვალსაზრისით ისეთ „უპერსპექტივო“ სიტყვებზე, როგორებიცაა:

„თევზი“, „მთა“, „თავი“...

დარგის სპეციალისტთა შეფასებით, მიღებული იქნა საყურადღებო შედეგები. გამოიკვეთა ტერმინების ეტიმოლოგიური წარმოშობის საკითხთა განხილვისას შემოთავაზებული მიდგომის პერსპექტიულობა, მაგრამ ნათელი გახდა ისიც, რომ საერთო სურათის შექმნა არამცთუ „ყოველი ენისთვის“, მხოლოდ ქართულისთვისაც კი, კომპიუტერის დახმარების გარეშე ვერ მოხერხდებოდა.

საჭირო გახდა, შეგვეშეშინებინა კონცეფცია, რომელსაც დაეფუძნებოდა შესაბამისი დანიშნულების კომპიუტერული სისტემა. გადავწყვიტეთ დასახული მიზნის მისაღწევად დავერდნობოდით არსებულ ორიენტირებულ მიდგომას.

არსი წარმოგვიდგება როგორც მნიშვნელოვანი, “დასაბამიდან არსებული” რაიმე ობიექტი, მოვლენა თუ ქმედება (მდინარე, მიწა, ქვა, ცა, მნათობი... მცენარე, ცხოველი, ადამიანი, სხეულის ნაწილი, მოძრაობა...

წვიმა, ელვა, რიტუალური წეს-ჩვეულება...). არსები მოდელირდება კლასების მეშვეობით, მათი რეალიზება კი ხდება კლასის ეგზემპლარებით. სწორედ, ეგზემპლარებში დაკონკრეტდება კლასში აღწერილი არსის თვისებები და საჭიროების შემთხვევაში მათთვის გამოიძახება შესაბამისი მეთოდები, რომლებიც შეაფასებენ ორ კონკრეტულ სიტყვას შორის ეტიმოლოგიური კავშირის არსებობაზე ჰიპოთეზის მართებულობის ხარისხს. ამასთან, დასაშვებია ეს სიტყვები განსხვავებულ ენებზე იყოს მოცემული.

კლასთან აგრეგატული მიმართებით შეიძლება დაკავშირდეს ქვეკლასები, ხოლო ასოციაციური მიმართებით – განსხვავებული სახის ცნებებიც.

აგრეგატული მიმართებისათვის ლინგვისტიკის დარგში მიღებულია შემდეგი დასახელებანი:

*სემანტიკური დავიწროება და სემანტიკური გაფართოება.*

მაგალითად, ვიკილექსიკონებში ცალკე არის გამოტანილი მოცემული სიტყვის სხვა ცნებებთან დასაკავშირებლად პუნქტი ჰიპოთეზის, რომლის მეშვეობითაც გავდივართ სემანტიკური გაფართოების ტერმინებზე.

ამრიგად, დასაწყისში კი კომპიუტერული სისტემის მონაცემთა ბაზის შესაბამის ცხრილებში განთავსებული იქნება არა მარტო რამდენიმე ათეულ ენაზე არსებული “საინტერესო” სიტყვების მარაგი, არამედ ის ცოდნაც, რომელიც აპრიორი ცნობილი იქნება ამ ეტაპზე სპეციალისტებისათვის, მაგალითად, ინფორმაცია ცნებებს შორის არსებული აგრეგატული და ასოციაციური მიმართებების შესახებ და სხვ.

ქვემოთ მოგვყავს მაგალითი, თუ როგორ გვესახება ჩვენ მიერ შემოთავაზებული ასოციაციური მიმართებებისა და მათი ძიებების მექანიზმის ავტომატიზების არსი:

დავუშვათ, მომხმარებელი (ლინგვისტი, ექსპერტი, არასპეციალისტი) გამოთქვამს ვარაუდს, რომ ტერმინი “კომლი” ეტიმოლოგიურ კავშირში უნდა იყოს ტერმინ “კვამლთან”. ეს ვარაუდი ლოგიკას ექვემდებარება, მაგრამ ცალსახად მისი სისწორის მტკიცება გამართლებულად ვერ ჩაითვლება. მართლაც, ლინგვისტებისათვის ცნობილია, თუ ხშირად როგორ აფერხებს ეტიმოლოგიების დადგენის საქმეს ე.წ. ხალხური ეტიმოლოგიები, რაც ძირითადად კონტამინაციის მოვლენით არის განპირობებული. სწორედ, ამ საქმეში არის უპირანი კომპიუტერის დახმარება – ჩვენ მიერ გაცემული ბრძანებით შესაძლებელი იქნება შესაბამისი დანიშნულების პროგრამული მოდულის გამოძახება, რომელიც შეეცდება მოცემულ და/ან სხვა ენებზე მოძებნოს ასეთივე კავშირის არსებობაზე “ეჭვმისატანი” სიტყვათა წყვილები. თუ აღნიშნული მოდული კარგად იქნება მოფიქრებული და რეალიზებული, ძიებამ უნდა მოგვცეს ასეთი შედეგები:

კერა – კერია, – , – , ...

ექსპერტები გაანალიზებენ კომპიუტერული სისტემის მიერ მოწოდებულ ვარიანტებს და თუ შეთანხმდებიან არსებული ჰიპოთეზის დიდი ალბათობით მართებულობაზე, კომპიუტერული სისტემის ცოდნის ბაზაში შეიტანენ შესაბამის ინფორმაციას.

მოცემული სიტყვისათვის სხვადასხვა ენებზე შესაძლო ეტიმოლოგიურ კავშირში მყოფი ტერმინების მოძიების გარდა, კომპიუტერულ სისტემას შეეძლება სხვა სახის მომსახურებების გაწევა. მოვიყვანოთ მაგალითი, თუ კიდევ როგორი სახის ძიებების ჩატარების დაკისრებას ვვარაუდობთ მისთვის:

ცნობილია, რომ აზიაში არსებობს მდინარე, რომლის ხუთმარცვლოვან დასახელებაში შემავალი თითოეული მარცვალი აღნიშნავს ცნება წყალს (მდინარეს) ამ რეგიონში ადრე და ამჟამად მცხოვრები ხალხების ენებზე.

მრავლისმეტყველი ფაქტია!

მოვანდინოთ ამ ინფორმაციისაგან ”ექსტრაქტის” გამოყოფა და ვაქციოთ იგი კომპიუტერული სისტემაში არსებული ცოდნის ბაზის ერთ-ერთ ელემენტად, რათა შემდგომში მასზე დაყრდნობით სისტემას დავაკლოთ შესაბამისი ძიებების ჩატარება.

მხოლოდ საქართველოში არსებული მდინარეების სახელწოდებების ანალიზის შედეგად მოსალოდნელია, კომპიუტერმა შემოგვთავაზოს მრავალი სამუშაო ჰიპოთეზა – მაგალითად, ზემოთ მოყვანილი მაგალითის ანალოგიურად, შესაძლებელია პროგრამამ გასაანალიზებლად მოგვაწოდოს შემდეგი სავარაუდო კომპოზიციები:

მატ-კვარი (მატ და კვირ ტერმინები თვით არიან ეტიმოლოგიურად “წყლის” შთამომავალი), ენ-კური (მდრ. ანკარა) და სხვ.

ცოდნის ბაზაში შესაძლებელი იქნება, ჩაიღოს ლინგვისტებისათვის ცნობილი მრავალი სხვა ფაქტის კომპიუტერთან დიალოგში გაანალიზების შედეგად ახლად ფორმირებული ცოდნის ელემენტი, მაგალითად ის დაკვირვება, რომ გრძელი მდინარეების დასახელებანი გაცილებით იშვიათად იცვლება და ზოგადად, უფრო არქაულია, ვიდრე მცირეთა სახელწოდებანი და სხვ.

დასასრულ, ქართული ენა წმინდა ლინგვისტიკური თვალსაზრისითაც უაღრესად საინტერესო ფენომენს წარმოადგენს. მისი მეშვეობით შესაძლებელია, ნაპოვნი იქნეს ის დაკარგული ჯაჭვის რგოლები, რომელთა არსებულებთან გამთლიანება უდავოდ შეუწყობს ხელს მსოფლიოში არსებული ენების საერთო წინაპრის – დღესდღეობით ჰიპოთეტურად მიჩნეული ბორეული ენის რეკონსტრუქციას, მით უფრო, თუკი ამ

საქმეში დავიხმარებთ დღევანდელი კომპიუტერების მძლავრ შესაძლებლობებზე დაფუძნებულ, მოხსენებაში შემოთავაზებულ კომპიუტერული ექსპერტული სისტემის კონცეფციას.

### 3. დასკვნა

შემუშავებულია სხვადასხვა ენებზე სიტყვების ეტიმოლოგიის ძიებისთვის კომპიუტერული სისტემის აგების კონცეფცია, რომელიც ეფუძნება არსზე ორიენტირებული დაპროგრამების პარადიგმას – სისტემის ბაზაში სხვადასხვა ენებზე დაფიქსირებულ სიტყვებს შორის მყარდება აგრეგატული და ასოციაციური კავშირები.

კომპიუტერული სისტემისათვის შემუშავებული კონცეფციის საფუძველია ავტორის მიერ სხვადასხვა მეთოდების გამოყენებით გაკეთებული დასკვნა:

თავდაპირველი ლექსიკური მარაგის შემადგენელი სიტყვების უმრავლესობისათვის ამოსავალი – ფესვური ცნება არის **წყალი**.

პროგრამა აფასებს მსგავსი უღერადობის სიტყვებს შორის ეტიმოლოგიური კავშირის შესაძლებლობას, ეძებს სიტყვათა წყვილებს შორის მოცემული ასოციაციური კავშირის მქონე ანალოგიურ წყვილებს ენათა სამყაროში და სხვ. საბოლოო სიტყვა კი ეკუთვნით ექსპერტებს.

ეტიმოლოგიური ძიებებისთვის განკუთვნილი მოდული შესაძლებელია დაემატოს ვიკიპედიას ერთ-ერთ კომპონენტად.

#### ლიტერატურა:

1. [http://en.wiktionary.org/wiki/Wiktionary:Main\\_Page](http://en.wiktionary.org/wiki/Wiktionary:Main_Page)
2. [http://fr.wiktionary.org/wiki/Wiktionnaire:Page\\_d%E2%80%99accueil](http://fr.wiktionary.org/wiki/Wiktionnaire:Page_d%E2%80%99accueil)
3. <http://de.wiktionary.org/wiki/Wiktionary:Hauptseite>
4. <http://ru.wiktionary.org/wiki/>
5. <http://ru.wiktionary.org/wiki/>
6. [http://gtu.ge/books/g\\_vinepadze\\_Qebai.pdf](http://gtu.ge/books/g_vinepadze_Qebai.pdf)
7. ღვინევაძე გ. არსზე ორიენტირებული ლინგვისტიკა. სამეცნ.კონფ. „ბუნებრივ ენათა დამუშავება. ქართული ენა და კომპიუტერული ტექნოლოგიები“. თბ., 2004.

### CONCEPT OF A COMPUTER SYSTEM OF ETYMOLOGICAL SEARCH

Gvinepadze Gela

Georgian Technical University

#### Summary

Ascertainment of the etymology of words is a complex linguistic task, especially when the original vocabulary is concerned. The investigations carried out by us resulted in the following: for the most part of the original words the key point is the basic concept which is believed to be WATER. According to this supposition, 'hopeless' sequence of the ways of formation of Georgian words has been ascertained. Certainly in order to extend the investigation, the utilization of the computer has become necessary. The conception concerning the automated system of the etymological search of words has been elaborated. The program estimates the possibilities of etymological linkage between the words of the similar phonation but at the same time the paper states that the concluding remarks are to be made by the experts.

T

Резюме

« »