

კონფერენცია

ბუნებრივ ენათა დამზადება

ქართული ენა და კომპიუტერული
ტექნოლოგიები



თბილისი

2003

საქართველოს პრეზიდენტთან არსებული
სახელმწიფო ენის მუდმივი სახელმწიფო კომისია

საქართველოს მეცნიერებათა აკადემიის არნ. ჩიქობავას
სახელობის ენათშეცნიერების ინსტიტუტი

ბუნებრივ ენათა დამუშავება ქართული ენა და კომპიუტერული ტექნოლოგიები

2003 წლის 16–18 ივნისი

კონფერენციის მასალები



„ქართული ენა“
თბილისი
2003

მუშაობის გეგმა

ორშაბათი, 16 ივნისი

- 10.15 გ. კვარაცხელია — შესავალი სიტყვა
10.30 მ. სუხიშვილი, ლ. ბერძენიშვილი, ი. ღარიბაშვილი, თ. ჩხერიშვილი — ინტერნეტის ინგლისურ-ქართული ლექსიკონის შესახებ
11.00 გ. ჩაჩანიძე — კომპიუტერული ლექსიკოგრაფია და ლექსიკონთა ტი-
პოლოგიის საკითხი
11.30 ი. ჩანტლაძე, ქ. მარგიანი, რ. იოსელიანი — სვანური ტექსტის სიმ-
ფონია-ლექსიკონი და სვანურ პუბლიკაციათა კომპიუტერულ ასახ-
ვასთან დაკავშირებული პრობლემები
12.00 დ. მელიქიშვილი — ქართული ზმნის ულლებადი ფორმების
კომპიუტერული მოდელირებისათვის

12.30-14.00 — შესვენება

- 14.00 კ. ფხაკაძე, გ. ჭანკვეტაძე, ლ. ტიბუა, მ. ივანიშვილი, ლ. ლეკიაშ-
ვილი, ე. სოსელია — ენობრივ-აზროვნებითი გაგებადობის ცნებისა
და მარტივი და ინტეგრალური გაგების ზმნური სიტყვების შესახებ
14.30 მ. ივანიშვილი, ლ. ლეკიაშვილი, ე. სოსელია, კ. ფხაკაძე, გ. ჭან-
კვეტაძე, ლ. ტიბუა — ლოგიკური, პროგრამული, ბუნებრივი ენები
— სემანტიკური სისტემის სამი განსხვავებული ტიპი
15.00 გ. ჩაჩანიძე, ლ. ბერიაშვილი — ელექტრონული კარტოთეკა და ქარ-
თული ენის განმარტებითი ლექსიკონის სადოკუმენტაციო ფონდი
რ. ბერულავა — სიტყვათშემოკლებათა სისტემის პროექტისათვის
კომპიუტერულ ტერმინოლოგიაში
16.00 ტ. ფუტკარაძე, ო. შენგელია — ქართველური დიალექტური ტექს-
ტების ელექტრონული ბაზა: პრობლემები, პერსპექტივები

საშუალებათი, 17 ოვნიხი

- 10.00 **თ. უთურგაიძე, თ. ვაშაკიძე, მ. მანჯგალაძე, ლ. ეზუგბაია, გ. ქობერიძე** — მრავლობითი რიცხვის ნულოვანი ალომორფები
- 10.30 **ჭ. ანთიძე, დ. მიშელაშვილი** — ინსტრუმენტული საშუალებები ქართული ტექსტების კომპიუტერული დამუშავებისათვის
- 11.00 **რ. ლანდია, თ. უთურგაიძე, ლ. ჩხაიძე, ი. გიბერტი** — ძველი ქართული ენის მანქანური ლექსიკონის ავტომატიზებული გენერაციის სისტემა (სალიტერატურო ძეგლთა ტექსტების ბაზაზე)
- 11.30 **გ. გურაბანიძე, მ. დუმბაძე** — „ქართლის ცხოვრების“ კომპიუტერული მოდელი
- 12.00 **ნ. ხუციშვილი** — საისტორიო დოკუმენტების კომპიუტერული დამუშავების საკითხები პ. კეკელიძის სახ. ხელნაწერთა ინსტიტუტში
- 12.30-14.00 — შესვენება
- 14.00 **ჭ. ანთიძე, ს. შენგელია** — ქართული წინადადების სინტაქსური და სემანტიკური სტრუქტურის წარმოდგენის შესახებ
- 14.30 **გ. გოგიჩაიშვილი, მ. ვეკუა, ლ. ჩხაიძე** — მარტივი წინადადების სინტაქსური ანალიზატორი შეზღუდული ქართული ენისათვის
- 15.00 **შ. დავითელაშვილი, ლ. ჩხაიძე** — საინფორმაციო სისტემა ნახური ენის ლექსიკონის ავტომატიზებული დაპროექტებისათვის
- 15.30 **ო. ქაპანაძე** — მრავალენოვანი ვალენტური ლექსიკონის არქიტექტურა ენის ტერმინოლოგიისათვის: ქართულ-რუსულ-ინგლისურ-გერმანული (GREG) ვალენტური ლექსიკონის ნიმუში

ოთხშაბათი, 18 ივნისი

- 10.00 თ. უთურგაძე, გ. ჩაჩანიძე, ლ. ჩხაიძე, ი. ჭანაშვილი, პ. მოირერი — ქართული ზმნის პარადიგმების კომპიუტერული ტაქსონომია
- 10.30 ქ. დათუკიშვილი, ნ. ლოლაძე, გ. ზაკალაშვილი — ქართული ენის მანქანური დამუშავება (მორფოლოგიის დონე)
- 11.00 ა. შადური, ი. შადური — აფხაზურ ზმნურ სიტყვაფორმათა სინთეზი
- 11.30 გ. კრავჩინკო, ლ. ჩხაიძე, ც. კვანტალიანი, ც. ჭანჭლავა — ქართული ზმნის ცოდნის ბაზის აქტუალიზაციისა და სრულყოფის ავტომატიზებული სისტემის შექმნის შესახებ
- 12.00 ლ. მარგველანი — ქართული სიტყვა-ფორმების მანქანური ანალიზური სისტემა

12.30-14.00 — შესვენება

- 14.00 გ. შერვაშიძე — ქართული ენა და საინფორმაციო ტექნოლოგიები
- 14.30 ჸ. ინგია — ენის მანქანური მოდელირება და პროგრამული ხელსაწყოები
- 15.00 ქ. ორკოშნელი — ასო-ნიშანთა სიხშირულ-სტატისტიკური გამოკვლევის მეთოდება და გამოყენების სფერო
- 15.30 ლ. ჩხაიძე, დ. ჭაში, ი. იაკობაშვილი — ქართულენოვან ტექსტებში არამკაფიო ძებნის მეთოდების დამუშავება და მათი გამოყენება ქართული პრესის ელექტრონული არქივის ინტერნეტის საძიებო სისტემაში

2.4. ანალიზის პროგრამა. ანალიზის პროგრამით ხდება იმ სიტყვა-ფორმათა ანალიზი, რომელიც სინთეზირდება. კერძოდ, პროგრამა შლის სიტყვაფორმებს მორფემებად, გამოყოფს ფუძეს, აძლევს კვალიფიკაციას თითო-ეულ მორფემას, განსაზღვრავს სახელის ბრუნვასა და რიცხვს, ზმნების პირს, რიცხვს, მწკრივსა და ა.შ.

ქართული ენის მორფოლოგიური პროცესორი იძლევა ქართული ენის მანქანური დამუშავების უამრავ შესაძებლობას. კერძოდ, პროცესორით შესაძლებელია ენის სიხშირული მახასიათებლების დადგენა და სტატისტიკური ანალიზი, ქართულენოვან ტექსტებში საჭირო გრამატიკული ფორმების მოძიება, სხვადასხვა ტიპის ლექსიკონების შედგენა და ა.შ.

სამუშაო შესრულდა ფონდ „ღია საზოგადოება — საქართველოს“ მიერ დაფინანსებული ორი პროექტის ფარგლებში.

3. ივანიშვილი, ლ. ლექიაშვილი, ე. სოსელია

საქართველოს მეცნიერებათა აკადემიის გ. წერეთლის სახ.
აღმოსავლეთმცოდნეობის ინსტიტუტი, ლოგიკისა და ენის ქართული ჯგუფი
ქ. ფხავაძე, გ. ჭანკვეტაძე, ლ. ტიბუა

თსუ ი. ვეკუას სახ. გამოყენებითი მათემატიკის ინსტიტუტი,
ლოგიკისა და ენის ქართული ჯგუფი

ლოგიკური, პროგრამული, ბუნებრივი ენები — სემანტიკური სისტემის სამი განსხვავებული ტიპი?

რ. ჰაუსერი გამოყოფს სამი პრინციპულად განსხვავებული ტიპის ენებს: ლოგიკურ ენებს, პროგრამულ ენებს და ბუნებრივ ენებს და შესაბამისად სამი განსხვავებული ტიპის სემანტიკას. იგი ამბობს: ლინგვისტიკაში სემანტიკა გრამატიკის შემადგენელი ნაწილია, რომელიც სინტაქსურად გაანალიზებული ბუნებრივი ენის ზედაპირული ფორმებისაგან წარმოქმნის მნიშვნელობის გამოხატულებებს. ანუ მისთვის სემანტიკა, მნიშვნელობა არაპირველადი, მეორადი გრამატიკული მოვლენაა, რომელიც უკვე მანამადე გრამატიკულად აგებული ენობრივი მასალის სინტაქსური ანალიზის შედეგად მიიღება — ჩვენ ამ თვალსაზრისს არ ვიზიარებთ. ყოველ შემთხვევაში, ბუნებრივ ქართულ ენობრივ სისტემაში, რომელმაც უფრო ბუნებრივად და თავისუფლად შეისისხლობორცა ფრეგესეული კომპოზიციურობის პრინციპი განსხვავებით მონტეგიუსეული მცდელობისაგან შეეტანა იგი ინგლისურ ენობრივ სისტემაში, გვარწმუნებს ჩვენ მიერ ზემოთ აღნიშნულში. ამგვარად, შეიძლება ითქვას, რომ ქართულ ბუნებრივ ენობრივ სისტემაში წარმოდგენილია ბუნებრივადვე არსებული ძირეული ფრეგესეული კომპოზიციური სემანტიკა. სწორედ ეს იყო საფუძველი ჩვენ მიერ უკვე გაკეთებული განაცხადისა ქართული ენის ტიპობრივ თანმთხვევაზე ფრეგესეული ტიპის ფორმალურ-მათემატიკურ ენებთან.

რ. ჰაუსერისათვის ყოველგვარი სემანტიკა არის ორდონიანი ბაზისური სტრუქტურა, რაც შედგება ენობრივი გამონათქვამებისა და მასთან დაკავშირებული სემანტიკური სტრუქტურებისაგან, რომელთა სინტაქსური გაანალიზება შესაძლებელია. ეს ორი დონე ერთმანეთს უკავშირდება განაწილების ალგორითმის საშუალებით:

I. ენის ზედაპირული დონე

სინტაქსურ-სემანტიკური
განაწილების ალგორითმი

↓↑

II. სემანტიკური დონე

თეორიულად, ერთი და იმავე ენისათვის სხვადასხვა სემანტიკური ინტერპრეტაცია შეიძლება განისაზღვროს სხვადასხვა განაწილების ალგორითმის გამოყენებით. პრაქტიკულად, სემანტიკის ყოველ ტიპს ახასიათებს მისი საკუთარი სინტაქსი, იმ მოთხოვნით, რომ მიღწეულ იქნეს სემანტიკური დონის ოპტიმალური კონტროლი ენის ზედაპირული ფორმების შეშვეობით. ამით რ. ჰაუსერი კიდევ ერთხელ ადასტურებს თვალსაზრისს ენის გრამატიკულ შრეში სემანტიკის მეორადულობაზე. ჩვენთვის უსემანტიკო ენობრივი ერთეული არ არსებობს ენობრივ-გრამატიკული სისტემის არცერთ დონეზე, გარდა ფონემატური საფეხურისა. ფონემები ენობრივი ნიშნების შექმნელი საშუალებებია. ამასთან ენობრივი ნიშანი გაგებულია როგორც მაკავშირებელი შუალედური რგოლი ნიშნის (ენობრივი ნიშნის) სამყაროსეულ მნიშვნელობასა და აზრით მნიშვნელობას შორის. ანუ ჩვენთვის ენობრივი ნიშანი სამდონიანი, საშრიანი განუყოფელი და ძირეული (ფუძემდებლური) სემანტიკური სტრუქტურაა.

რ. ჰაუსერის მიხედვით არსებობს სემანტიკური სისტემის საში განსხვავებული ტიპი:

1. ლოგიკური ენები — სათავეს იღებენ ფილოსოფიიდან (არისტოტელე). ლოგიკური ენების სრულ გამოსახულებებს გამონათქვამები ეწოდება. მათი დანიშნულებაა განსაზღვრონ ჭეშმარიტების მიმართება ნებისმიერ წინადადებასთან ნებისმიერ მოდელში. მათ აქვთ მეტაენაზე დაფუძნებული სემანტიკა, რადგან ორ დონეს შორის კორელაცია ემყარება მეტაენის განსაზღვრებას.

2. პროგრამული ენები — შექმნილია პრაქტიკული საჭიროებით, რათა გამარტივდეს კომპიუტერთან (აბსტრაქტულ მექანიზმთან) ურთიერთობა. პროგრამული ენების გამოსახულებებს ბრძანებები ეწოდება, რომლებიც შესაძლებელია გაერთიანდეს რთულ პროგრამებად. მათ აქვთ პროცედურული სემანტიკა, რადგან კორელაცია სინტაქსურ და სემანტიკურ დონეებს შორის ემყარება შესრულების პრინციპს, ანუ ბრძანებათა ოპერაციულ რეალიზაციას აბსტრაქტული მექანიზმის მეშვეობით, რომელიც ჩვეულებრივ ხორციელდება ელექტრონულად.

ჩვენთვის პროგრამული ენის ნებისმიერი ძირეული, განუყოფელი ნიშანი, ასევე ჩვეულებრივი სამდონიანი, საშრიანი სტრუქტურაა. რაც შეეხება თავად კომპიუტერს (აბსტრაქტულ მექანიზმს), იმ შემთხვევაში, თუ მას შემეცნებითი უნარით არ აღვეურვავთ, მაშინ ეს ნიშანი ლილაკის მსგავსი ორგონიანი სემანტიკური სტრუქტურა იქნება, ხოლო თუ მას შემეცნებითი უნარებით დავინახავთ (რაც უფრო სწორი ხედვაა), მაშინ იქაც სამდონიან ნიშნით სტრუქტურასთან გვექნება საქმე. რაც შეეხება ნიშნის სემანტიკურ პროცედურულ-ბრძანებით ბუნებას, ეს მისი (აბსტრაქტული მექანიზმის) შემეცნებითი სპეციფიკაა და ჩვენთვის იგი ამ შემთხვევაში ენობრივად არაარსებითია.

თანამედროვე ლოგიკურ ფორმალურ-მათემატიკურ ენებს დღეს არსებითად აღარ განასხვავებენ კომპიუტერული პროგრამული ენებისაგან. იმპერატიული და არაიმპერატიული პროგრამირების განკოფა და ამ ტერმინების გააზრება ემყარება სწორედ ე. წ. ხელოვნური ფორმალურ-მათემატიკური ენების ფარგლებში განვითარებული ლოგიკური პროგრამირების იდეებს. ჩვენთვის ლოგიკური ენის ნიშანიც სამდონიანი, საშრიანი განუყოფელი სტრუქ-

ტურაა და აქედან გამომდინარე, სილრმისეული ძირეული ნიშნობრივი ბუნებებით ეს განსხვავებული ტიპის ენები ერთმანეთისაგან არ განიჩინევიან.

რ. ჰაუსერთან: ბუნებრივი ენები ვითარდება ბუნებრივად სამეტყველო კომუნიკაციებში, ამიტომ სემანტიკურად ინტერპრეტირებადი სისტემებიდან ისინი ყველაზე ძლიერი და ძნელად ახსნადნი არიან. ბუნებრივი ენის გამოსახულებებს ენის ზედაპირული ფორმები ეწოდება. ენის ზედაპირულ ფორმებთან დაკავშირებული სემანტიკური წარმოდგენები გამოიყვანება ბუნებრივი კომუნიკაციების ზოგადი პრიციპებით, რადგან მნიშვნელობის ზედაპირულ ფორმებს მსგავსი კონკრეტული მანიფესტაცია არა აქვთ. ბუნებრივ ენებს აქვთ პირობითად დაფუძნებული სემანტიკა და სიტყვათა ზედაპირულ ფორმებს (ტიპებს) მნიშვნელობები შეთანხმების საფუძველზე უვითარდებათ. ვეთანხმებით ასეთ წარმოდგენას იმ ნაწილში, რომ ბუნებრივი ენა ბუნებრივადვე განვითარებადი ენაა, რისი საშუალებითაც კომუნიკაციის პროცესში არსებითად ვახერხებთ ნებისმიერი აზრის ენობრივად გამოხატვას, ნებისმიერი აზრის ენობრივ ფორმებში მოქმედვას. ამასთან, ჩვენი თვალსაზრისით, მსმენელის ინტერესებიდან გამომდინარე, შესაძლებლობა გვეძლევა ენობრივად გაფორმებული ფრაზიდან აზროვნებითი გაგების გამოტანისა, ანუ აზროვნებითი აღქმის (წარმოდგენის) შექმნისა, იმ საფუძველზე, რასაც კომუნიკაციის პროცესში ტექსტთან ერთად მოწოდებული ტექსტური მასალის საინტერპრეტაციო-პირობითი გარემო წარმოადგენს. ეს საბოლოო ჯამში იგივე თვალსაზრისია, რასაც რ. ჰაუსერი ავითარებს. ოღონდ უფრო გასაგები ხდება, თუ რატომ არ გააჩნიათ ზედაპირულ ფორმებს კონკრეტული მანიფესტანტები მთლიანი ენობრივი სისტემის გასწვრივ და როგორ იტვირთებიან ისინი კონკრეტული მანიშნებლობითი ფუნქციებით ყოველ კონკრეტულ საკომუნიკაციო შემთხვევაში.

ამგვარად, ჩვენი თვალსაზრისით, არ არსებობს პრინციპული განსხვავება ლოგიკურ, პროგრამულ და ბუნებრივ ენებს შორის. თუმცა გასაგებია, რომ ბუნებრივი ენობრივი სისტემა, გამომდინარე მისი მატარებელი ერთეულის ფართო შემეცნებითი უნარებიდან, ბევრად უფრო რთულად აღსაწერი და ძნელად გასაცნობიერებელი სისტემაა, ვიდრე დანარჩენი ორი. გასაგებია ისიც, რომ ფრეგესეული ტიპის სისტემებად ცნობილი ფორმალურ-მათემატიკური ენობრივი სისტემები მხოლოდ იმ შემთხვევაში შეიძლება იყოს გააზრებული ბუნებრივი ენობრივი სისტემის ეკვივალენტური ძალის სისტემებად, თუ მათ ფორმალურადვე აღვტურვავთ ფორმალურად განვითარებადი რეინტერპრეტაციული უნარით. ცხადია, რომ თუ რაიმეს თქმა შეიძლება, შეიძლება მისი თქმა ბუნებრივი ენითაც. ანუ, თუ რაიმეს თქმას ვერ ვახერხებთ ბუნებრივი ენით, მაშინ მის თქმას ვერ მოვახერხებთ საერთოდ. გასაგებია ისიც, რომ ე. წ. ხელოვნური ფორმალურ-მათემატიკური სისტემა მხოლოდ იმ შემთხვევაში იქნება ბუნებრივი ენის ეკვივალენტური ენობრივი სისტემა გამომხატველობითი თვალსაზრისით, თუ ამ ფორმალურ-მათემატიკურ სისტემაში შევძლებთ ბუნებრივი ენობრივი სისტემის შემადგენელი თითოეული სიტყვის თქმას (აღწერას) ამავე ფორმალურ-მათემატიკური სისტემის ენობრივი საშუალებებით. ამდენად, ბუნებრივი ენობრივი სისტემის შემადგენელი სიტყვების პირდაპირ ფორმალურ-ლოგიკურ აღწერას ანუ მათ თქმას პრინციპული და გადამწყვეტი როლი ენიჭება როგორც ბუნებრივ ენობრივ სისტემათა ფორმალიზაციისა და მათემატიზაციის ინტერესებიდან გამომდინარე, ისე იმ ჩვენეული თვალსაზრისის დასასაბუთებლად, რასაც ლოგიკური, პროგრამული და ბუნებრივი ენების ზოგადი სემანტიკური ერთტიპობრიობა წარმოადგენს.

წინების ლექსიკა, ბავშვის ლექსიკონი, თამაშობები, ნათესაობის ლექსიკა, ღვთისმსახურება, ტანისამოსი, საჭმელ-სასმელი, ზომა-წონა, ხელობა, მესაქონლეობა, მეფრინველეობა, მეფუტკრეობა, მეაბრეშუმეობა, მესიმინდეობა, სამკურნალო მცენარეები, მეწისქვილეობა, მეჩაიეობა, მეციტრუსეობა, მეთამბაქოეობა, მეთევზეობა და მონადირეობის ლექსიკა...

ინსტიტუტის მომძიებელთა ჯგუფების შეირ 1993-1999 წლებში სულ ჩაწერილ იქნა 460-ზე მეტი აუდიოკასეტი (23000-ზე მეტი ხელნაწერი თაბანის ფურცელი), 800-ზე მეტი საგამომცემლო ფორმის მასალა. შასალა მოძიებული და ჩაწერილ იქნა შემდეგ სოფლებში: ორპირში, კურსებში (ტყიბულის რაიონი), მათხოვში, კუხში (ხონის რაიონი), ლენჯერში, ლატალში (მესტიის რაიონი), ქვალონში (ხობის რაიონი), დიდი ჭყონში (მარტვილის რაიონი) და ქუთაისში.

თემატურად დალაგებული და კომენტარებით აღჭურვილი მთელი მასალა შენახულია ელექტრონული ვერსიის სახით და მზადდება გლობალურ ელექტრონულ ქსელში გამოსაქვეყნებლად. მომძიებელთა ჯგუფების შეირ ჩაწერილი აუდიომასალა უკვე იწერება ლაზერულ დისკებზე და მზადდება განუსაზღვრული ვადით შესანახად.

ქსუ დიალექტოლოგიის ინსტიტუტთან მუშაობს სპეციალური ჯგუფი ციფრული ტექნიკით დიალექტური მასალის უშუალოდ ტექსტური ჩანაწერისა და საძიებო სისტემების შესაქმნელად.

ქ. ფხაკაძე, გ. ჭანკვეტაძე, ლ. ტიბუა

თსუ ი. ვეკუას სახ. გამოყენებითი მათემატიკის ინსტიტუტი,
ენისა და ლოგიკის ქართული ჯგუფი

გ. ივანიშვილი, ლ. ლეკიაშვილი, ე. სოსელია

საქართველოს მეცნიერებათა აკადემიის აკდ. გ. წერეთლის სახ. აღმოსავლეთ-
მცოდნეობის ინსტიტუტი, ენისა და ლოგიკის ქართული ჯგუფი

ენობრივ-აზროვნებითი გაგებადობის ცნებისა და მარტივი და ინტეგრალური გაგების ზმნური სიტყვების შესახებ

ჩვენ ვმოძრაობთ პასკალის იმ აზრის მიხედვით, რომელიც შემდეგი სიტყვებითაა გამოთქმული: „რა თქმა უნდა, არსებობს არაგანსაზღვრებადი სიტყვები, და ბუნებას რომ ჩვენთვის არ მოეცა მათი პირველსაწყისი გაგებანი, მაშინ ყველა დანარჩენი სიტყვაც და ფრაზაც ჩვენთვის ასევე გაუგებარი იქნებოდა!“ — პასკალი არა მხოლოდ ენას, არამედ ამ ენით რეალიზებულ ენობრივ-აზროვნებით უნარსაც ბუნებრივ ფენომენად, ბუნებრივ მოვლენად მიიჩნევდა. პასკალი არ გახლდათ ამ თვალსაზრისების პირველმქადაგებელი. იგი ძველი სამყაროდან მომდინარე ბიბლიური და, მასთან შედარებით უფრო ახალი, ძველი ბერძნული ლინგვისტურ-ლოგიკური მსოფლმხედველობით იყო აღჭურვილი, ამასთანავე, ლაიბნიცისა და დეკარტის ეპოქის შემადგენელი, ამიტომაც მრავალჯერ დაყენებულ კითხვებზე ერთგვარად დაზუსტებულ, ჩვენთვის უფრო ადვილად გასაგებ პასუხებს იძლევა. ჩვენ არ ვჩერდებით პასკალთან, რომელიც ვეუბიცყასეულ ხედვებს არსებითად წინამძღოლობს და ამჯერად არც მისეული (ვეუბიცყასეული) თანდაყოლილი კონცეპტების ცნების ხომსკისეულ ხედვებთან ნაჯვარობაზე ვაკეთებთ აქცენტს. არც მონტეგიუს ფრაზით ვიზლუდებით, რომელიც თვლიდა, რომ არ არის არანაირი არსებითი

განსხვავება ხელოვნურ ფორმალურ მათემატიკურ ენებსა და ბუნებრივ ენობრივ სისტემებს შორის. ჩვენ უფრო ღრმად მივდივართ: ჩვენ ვთვლით, რომ ბუნებრივი ენობრივი სისტემა არა ერთადერთი და არა პირველადი (არა ძირეული) ბუნებრივი ენობრივი სისტემაა და რომ ენობრივ სისტემაში არაგანსაზღვრებადი სიტყვების ბუნებით მოცემულ გაგებათა ერთი ნაწილი რეალიზდება ბუნებრივ ენობრივ სისტემასთან შედარებით უფრო სილრმისეულ, ვუწოდოთ ასე, ქვეცნობიერ ბუნებრივ-ენობრივ სისტემურ შრეში. ამგვარად: ბუნებრივი ენობრივი სისტემა ჩვენთვის არა ერთადერთი ბუნებრივი ენობრივი სისტემაა. ამდენად, მას ჩვენ ცნობიერ ბუნებრივ ენობრივ სისტემას ვუწოდებთ და ამგვარ ბუნებრივ ენობრივ სისტემებთან ერთად განვიხილავთ ჩვენ მიერ ქვეცნობიერად წოდებულ ბუნებრივ ენობრივ სისტემებსაც, რომლებსაც ჩვენი თვალსაზრისით შიდაადამიანისეულ ბუნებრივ ენობრივ სისტემათა შორის უფრო ძირეული ადგილი უჭირავთ, ვიდრე წინათ ხსენებულთ. ჩვენი თვალთახედვით ფრეგესეული, დღეს ხელოვნურ, ფორმალურ-მათემატიკურ ენებად წოდებული, ენათა კლასი სწორედ შიდაადამიანისეულ ქვეცნობიერ ბუნებრივ ენობრივ სისტემათა შემადგენელია და ის ტიპობრივი სიმარტივე, რაც ამგვარ ლოგიკურ-მათემატიკურ ენებს ბუნებრივი საურთიერთობო-საშეტყველო ენებისაგან განასხვავებს, არსებითი გაძყოფი და მაკლასიფიცირებელი პარამეტრია ცნობიერი და ქვეცნობიერი ბუნებრივი ენობრივი სისტემებისა.

ენობრივ-აზროვნებითი გაგებადობის ცნების გააზრებისას ჩვენ არსებითად ვეურდნობით ზემოთ გამოკვეთილ ენობრივ მსოფლხედვას და ცნობიერ ბუნებრივ ენობრივ სისტემაში არაგანსაზღვრებადი სიტყვების ე.წ. თანდაყოლილ ბუნებრივ გაგებებს ვშლით უფრო სილრმისეულ, ქვეცნობიერ ფრეგესეულ ფორმალურ-ლოგიკურ ბუნებრივ ენობრივ არეში ლოგიკურ-აქსიომატური საშუალებებით იმ უნივერსალურ სიმრავლურ-თეორიულ სემანტიკურ გარემოში, რომლის საფუძველზეც შონტეგიუ აყალიბებდა მისეულ თეზად წოდებულ ზემოთ მოტანილ თვალსაზრისს. ამასთან, მაკავშირებელ ხიდად ცნობიერი ბუნებრივი ენობრივი სისტემის არაგანსაზღვრებადი სიტყვებით აგებულ სისტემასა და ქვეცნობიერ ფორმალურ-მათემატიკურ ენობრივ სისტემას შორის ჩვენ ძირითადად ვიყენებთ პროფ. შ. ფხაკაძისეულ ფორმალურ აღნიშვნათა თეორიას, რომელიც ფრეგესეულ არაგანვითარებად ფორმალიზმს თვისებრივად ახალ ფორმალურად განვითარებადი ფორმალიზმის სახით აყალიბებს. გარდა ამისა, ჩვენთვის არსებითია თავად ცნობიერი ბუნებრივი ენობრივი სისტემის მკაცრი ბუნებრივ-სემანტიკური გადააზრება, ენობრივი სისტემისათვის დამახასიათებელი ბუნებრივად არსებული და ფორმალურად გატარებული მორფო-სინტაქსური წესრიგის განმსაზღვრელი საშუალებების არაორაზროვანი (არაწინააღმდეგობრივი) ფუნქციონალური ამოხსნა. ამ მხრივ განსაკუთრებულ სირთულეს ქმნის ქართული ზმნური სიტყვების მტკიცე და რთული მორფო-სინტაქსური კონსტრუქცია. ამასთან დაკავშირებით გამოვთქვამთ შემდეგ ფუძემდებლურ მოსაზრებებს: ზმნური სიტყვათა კლასი ქართულში ბუნებრივად იყოფა მარტივი და ინტეგრალური გაგების ზმნურ სიტყვათა კლასებად. ამასთან, ინტეგრალური გაგების ზმნური სიტყვები მარტივი გაგების ზმნურ სიტყვათა საფუძველზე განსაზღვრებადი კონსტრუქციებია, რომელთა მეშვეობითაც ინტეგრალური სიტყვათა გაგების საკითხი ბუნებრივად და მკაცრად-ფორმალურად დაიყვანება მარტივი გაგების ზმნურ სიტყვათა გაგებაზე. აქაც, მკაცრი განსაზღვრებების აგებისას ვისარგებლებთ ფორმალური აღნიშვნათა თეორიით, რომელიც დასაბუთებულად უზრუნველყოფს ფორმალიზმი-

სათვის დამახასიათებელი ჩასმის ფორმალური ოპერაციით სარგებლობის თავისუფალ და არაწინააღმდეგობრივ გარეშემოს.

ა. შადური, ი. შადური

ივ. ჯავახიშვილის სახ. თბილისის სახელმწიფო უნივერსიტეტი

აფხაზურ ზმნურ სიტყვაფორმათა სინთეზი

მოდელი მიზნად ისახავს სიტყვაფორმათა ძალიან დიდი რაოდენობის წარმოშობას მაქსიმალურად შარტივი გზით, ე.ი. შეძლებისდაგვარად მცირე-რიცხოვანი წესების საშუალებით.

წარმოდგენილი მოდელი შედგება სამი ნაწილისაგან: ფუძეთა ლექსიკო-ნისაგან, აფიქსთა რანგობრივი ცხრილებისა და წესების სისტემისაგან.

ფუძეთა ლექსიკონი მოცემულ ეტაპზე წარმოდგენილია ფრაგმენტის სახით, თუმცა მასში ასახულია ფუძეების ყველა ძირითადი ტიპი. რთული ფუძე-ები დაყოფილია შემადგენლებად. თითოეულ ფუძეს ლექსიკონში მინიჭებული აქვს რამდენიმე მახასიათებელი. სახელდობრ, მითითებულია ფუძის დინამიკუ-რობა-სტატიკურობა, ვალენტობა, გარდამავლობა-გარდაუვალობა, ბმულობა-არაბმულობა, ინვერსიულობა, ოდენ ლექსიკური ფუძით „სათავისო“ მნიშვნე-ლობის გამოხატვის შესაძლებლობა; ალრიცხულია ფუძის ორ ლექსემად და-შლის შემთხვევები და ზოგიერთი სხვა სპეციფიკური ნიშანი.

აფიქსთა რანგები პირობითად დაყოფილია ორ ჯგუფად — ფლექსიური და დერივაციული აფიქსების რანგებად. ამგვარი დაყოფა აადვილებს მორფო-ლოგიური წესების ბლოკში იმ შეზღუდვათა გათვალისწინებას, რომლებიც დაედება ლექსიკური ფუძეებისა და დერივაციული პრეფიქსების თანახმარე-ბას.

წესების სისტემა შედგება ბლოკებისაგან. ამჟამად დამუშავებულია მო-რფოლოგიური ანუ რანგებისა და ქვერანგების არჩევის წესები და მორფონო-ლოგიური წესები, რომლებითაც მორფემული მიმდევრობები გარდაიქმნება მორფულ მიმდევრობებად. ძირითად წესებს დართული აქვს რამდენიმე ფა-კულტატიური და აუცილებელი ტრანსფორმაცია. მათი საშუალებით შესაძ-ლებელია „არასტანდარტულ“ სტრუქტურთა წარმოშობა, რის შედეგადაც იზ-რდება სინთეზირებულ ფორმათა სიმრავლე.

მოდელი რეალიზებულია კომპიუტერული პროგრამის სახით, რაც შე-საძლებლობას გვაძლევს ბაზაში შეტანილი მორფემებისაგან შევადგინოთ ზმნურ სიტყვაფორმათა დიდი რაოდენობა. პროგრამა დაწერილია ენაზე C++.

წარმოდგენილი მოდელით წარმოიშობა გრამატიკულად სწორი მიმდევ-რობები. ზოგიერთი მათგანი ენობრივ უზუსში იშვიათად გვხვდება, მაგრამ გა-საგებია ამ ენაზე, როგორც მშობლიურ ენაზე, მოლაპარაკეთათვის.

მოდელს აქვს როგორც თეორიული, ისე პრაქტიკული მნიშვნელობა. მი-სი შესაძლო გამოყენების სფეროა სწავლების პრაქტიკა, აგრეთვე აფხაზური ენის ნორმატიული გრამატიკებისა და ლექსიკონების სრულყოფაზე მუშაობა. ანალიზის მიზნებისათვის მოდელის გამოსაყენებლად საკმარისია მისი მცირე-დი მოდიფიკაცია.

7) subject of the text (narration).

(Although there are languages with other successions or other forms of steps).

To study each object and to complete this information the recognition method is used. For recognition in Georgian these main principles are used:

for letter - comparison of the position of parts of a letter (lines, circles, arcs) with the database of positions;

for sound - comparison of dynamic, harmonic, spectrum characteristics with their database;

for word - morphological rules and database (dictionary);

for sentence - semantic rules, rules database and database of word connection in meaning.

Part of this system is realized in Georgian software: spell checker, translation into/from English, optical recognition, speech recognition, search engine... More than 50 software tools have been created to solve this tasks. The tools were used as in the solution of the task proper, so in the creation of databases and testing.

M. Ivanishvili, L. Lekashvili, E. Soseli

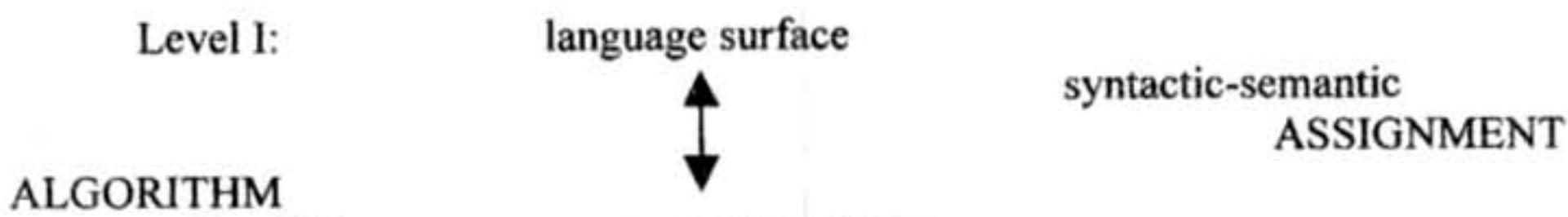
G. Tsereteli Institute of Oriental Studies, Georgian Group of Logic and Language

K. Pkhakadze, G. Chankvetadze, L. Tibua

I. Vekua Institute of Applied Mathematics, Georgian Group of Logic and Language

LOGICAL, PROGRAMMING, NATURAL LANGUAGES – DO THEY REPRESENT THREE DIFFERENT TYPES OF THE SEMANTIC SYSTEM?

R. Hausser distinguishes between the languages the three principally different types – Logical Languages, Programming Languages and Natural Languages and hence differentiates three different types of semantics: "In linguistics, semantics is a component of grammar which derives representations of meaning from syntactically analyzed natural surfaces." In other words, he regards semantics or meaning as a non-primary, secondary grammatical phenomenon, obtained via syntactic analysis of (previously) grammatically constructed linguistic material. We do not share this standpoint. We consider semantics a non-secondary, basic linguistic-mental phenomenon – at least in the natural Georgian language system which rather naturally and easily adopts Frege's principle of compositionality unlike Montague's attempt to put it in English Language system, we have possibility of understanding three existing Frege's type simple compositionality semantics. Hence we may assume that in the natural Georgian language system is present the basic compositional semantics of Frege – also naturally existent as such. On this very ground we have made our statement of the typological coincidence about the Georgian language with the formal-mathematical languages of Frege's type. R. Hausser regards any types of semantics as a "basic two-level structure consisting of syntactically analyzed language expressions and associated semantic structures. The two levels are systematically related by means of an assignment algorithm:



ALGORITHM

Level II:

language surface

syntactic-semantic

ASSIGNMENT

In theory, different semantic interpretations may be defined for one and the same language, using different assignment algorithms. In practice, however, each type of semantics has its own characteristic syntax in order to achieve optional control of the semantic level via the combination of language surfaces." By that R. Hausser once more confirms the view on the semantics as a secondary phenomenon. We believe that any linguistic unit of any grammatical level, except the phonemic one, has its semantics. Phonemes are the constructors of the linguistic sign and any sign is regarded as the connecting link between the universal meaning and mental meaning. In other words, we interpret the linguistic sign as an indivisible and asic (fundamental) semantic structure with three levels, with three strata.

According to R. Hausser, "there are three different types of semantic systems:

1. Logical languages. They originated in philosophy (Aristotle). Their complete expressions are called propositions. They are designed to determine the truth value of arbitrary propositions relative to arbitrary models. They have a metalinguage-based semantics because the correlation between the two levels is based on a metalinguage definition.

2. Programming languages. They are motivated by a practical need to simplify the interaction with computers and the design of software. Their expressions are called commands, which may be combined into complex programs. They have a procedural semantics because the correlation between the levels of syntax and semantics is based on the principle of execution, i.e., the operational realization of commands on an abstract machine which is usually implemented electrically." Any basic, indivisible sign of the programming language to us is the usual structure with three levels(strata), but for a computer (the abstract machine), in case of considering it as the one without consciousness, the sign will turn out the structure with two levels – like a button, but if we attribute consciousness to it (the more proper view on an abstract machine), then here too, we shall have the sign with three levels. And concerning with it the feature of procedural semantics of the sign-commands is cognitive characteristic (specific) of abstract machine and is of no relevance in this case.

Contemporary logical formal-mathematical languages are not principally distinguished from the programming languages. The differentiation of the imperative and non-imperative programming and the comprehension of these terms are based on the very concepts of the logical programming that were evolved in the limits of the artificial formal-mathematical languages. We regard the sign of the logical language as an indivisible structure with three levels or strata and hence, these languages of different types do not differ from each other by the fundamental nature of their signs.

Further, according to R. Hausser, there are "3. natural languages: They evolve naturally in their speech communities and are the most powerful and least understood of all semantically interpreted systems. Their expressions are called surfaces. The associated semantic representations have to be deduced via the general principles of natural communication because the meaning has no concrete

manifestation like the surfaces. Natural languages have a convention-based semantics because the word surfaces (types) have their meaning assigned by means of conventions. "We agree with such standpoint in the sense that a natural language is a naturally developed language by means of which in the communicating process we basically manage to express verbally, formulate linguistically any idea, any thought. At the same time, in our opinion, we, as the listeners are able to deduce the mental understanding from the linguistically formed phrase or create the mental conception according to the interpretational-conventional environment sent by the text. In the end it is the same conception that the one developed by R. Hausser. But now we can more easily explain why the surface forms have no concrete manifestations in the language system and how they obtain such concrete meanings in each particular communicating act.

Hence, in our opinion, there is no principal difference between logical, programming and natural languages, yet, resulting from the broad perceptual abilities of the users of the natural language system, it is much more difficult to describe and comprehend it than the languages of the other two types. It is evident as well that the formal-mathematical languages, the so-called systems of Frege's type, can be regarded as the systems of equal force to natural language system only in case if we formally endow them with the reinterpretation and formally developing abilities. Evidently, if it is possible to say something then it is also possible to say it via natural language, or, in other words, if it is impossible to say something by means of natural language, we would not be able also to say it in any other way. Apparently, the so-called formal-mathematical system may be considered equal with the natural language system from the expressive point of view only in that case if we can say (describe, express) every word of the natural language system in this formal-mathematical system only by means of this formal-mathematical language system.

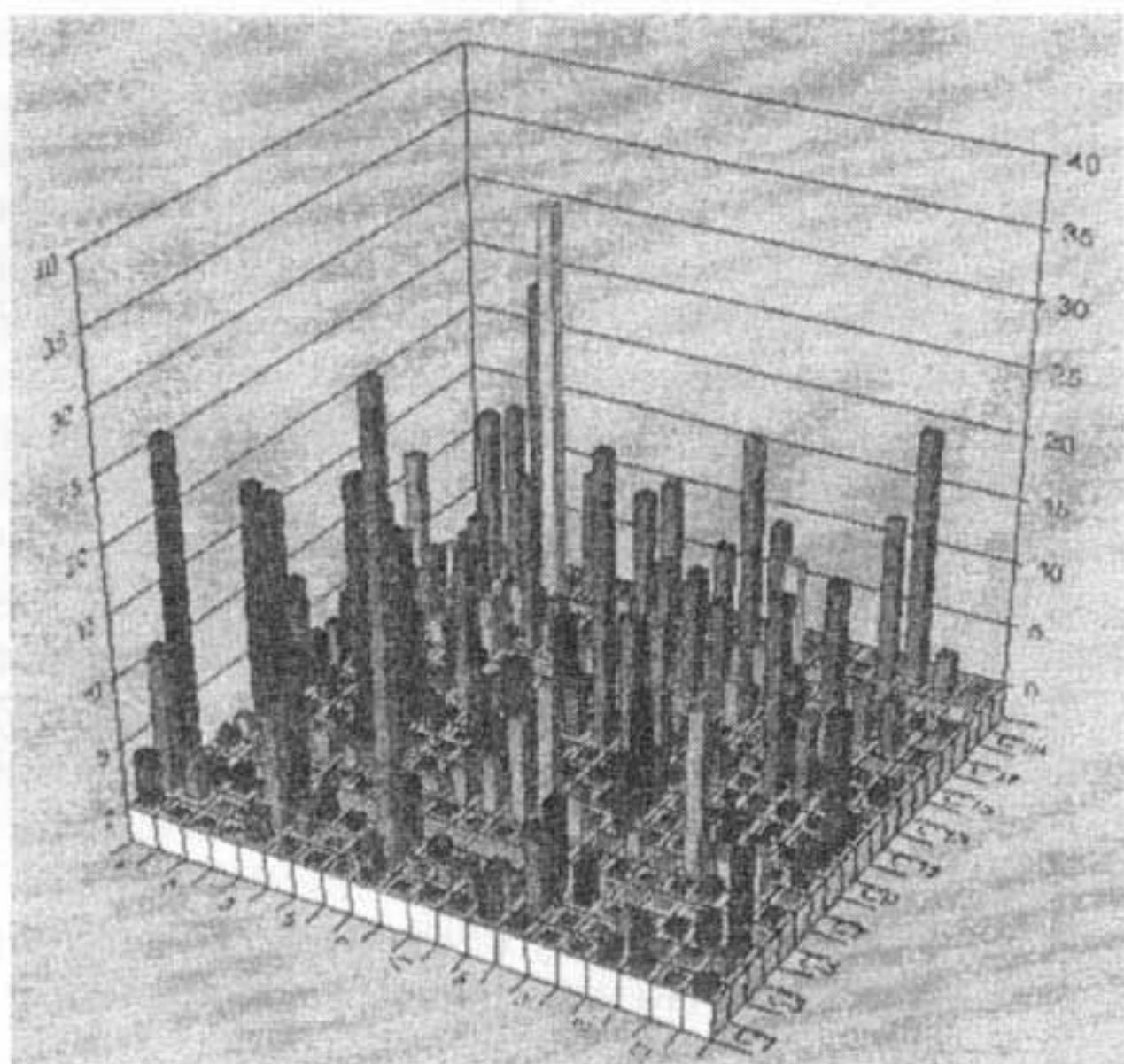
Hence, in view of formalization and mathematization of the natural language systems, the direct formal-logical description or, in other words, the expression of the natural language words via mathematical formal language proves to be of the top-priority importance.

O. Kapanadze
Tbilisi State University

A FRAMEWORK FOR MULTILINGUAL GEORGIAN, RUSSIAN, ENGLISH, GERMAN (GREG) VALENCY LEXICON

A *multilingual valency lexicon (MVL)* is a resource containing details of the words of more than one language. The details it contains are word valencies, for various situations in which the word is used, the lexicon characterises its complements (if any) both syntactically and semantically.

The least interesting kind of MVL would simply list, for each word, its different uses and the associated complements without describing the relations between corresponding words in different languages, or different uses of the same



Digraph letters laid on X and Y axis, and frequency on Z axis only digraphs with 15 letters that have the highest frequency are shown there. Also we can show deviation data via the same chart. This way it would be possible to learn easily an authors "style" and visually recognize an author. Another direction is the identification of a language or dialect automatically, according to the monograph statistical data.

K. Pkhakadze, G. Chankvetadze, L. Tibua

I. Vekua Institute of Applied Mathematics, Georgian Group of Logic and Language

M. Ivanishvili, L. Lekhiashvili, E. Soselia

G. Tsereteli Institute of Oriental Studies, Georgian Group of Logic and Language

ON THE CONCEPT OF MENTAL-LINGUISTIC UNDERSTANDABILITY AND ABOUT THE VERBS OF SIMPLE AND INTEGRATED UNDERSTANDING

We follow Pascal. He says: Evidently, there are words that cannot be defined and if the nature has not given us their initial concepts then all the other words and phrases would not be understandable to us. – Pascal regarded not only the linguistic capacity but also the linguistic-mental abilities as a phenomenon of Nature. Pascal was not the first propagator of these conceptions. He had absorbed the linguistic-logical viewpoints of the Bible and ancient Greece and as a child of the epoch of Leibniz and Decartes gives precise, easily understandable answers to the repeatedly arising questions. We do not dwell on Pascal's conception that actually precedes Verzbicka's approaches nor accentuate on the points of intersection between Verzbicka's innate concepts and Chomsky's attitudes; nor do we confine our

approach to Montague's viewpoint – the latter did not see any essential difference between the formal mathematical and natural languages. We go further: we believe that natural language system is not the only and initial (fundamental) natural language system and also, we believe that the intrinsic words (concepts) that cannot be defined in the language system are actualized in the more profound (compared with the natural language system) so – called subconscious stratum of the natural-language system. Hence, we do not regard the natural language system as the only one in-humanity natural language system. We call this natural language system a conscious natural language system and we think that subconscious language system is more intrinsic (profound) natural language system compared with the conscious natural language system. We are considering also so-called unconscious natural language stratum in-humanity language space but this is not now under consideration, though we can say that it (unlike the above-specifieds) occupies the most fundamental position among the intrinsic natural language systems. According to our conception the class of Frege's type languages (in modern definition the artificial, formal-mathematical languages) represents the intrinsic subconscious natural language systems and the simplicity of the typology that distinguishes such logical-mathematical languages from the natural communicational ones is the principle classificatory parameter of the conscious and subconscious natural language systems.

Considering the concept of linguistic-mental understandability we actually follow the above-specified linguistic viewpoint. So the words exclusively undefinable in the conscious natural language system are understood in the Frege's type subconscious language stratum by means of axiomatization in that set-theoretical area where Montague had developed his above-mentioned conception. At the same time we will mainly use the Formal Notation Theory by prof. Sh. Pkhakadze (the theory that re-forms Frege's non-extensible formalism into the formally extensible one) as a bridge between those words (concepts) of the conscious natural language system which cannot be defined by tools of this natural language and the Frege's type subconscious formal-mathematical language system. Besides, the strict natural-semantic understanding of the conscious natural language system seems a rather important task as well as the unambiguous (non-contradictory) functional solution of the marking mechanism and the nature of the morpho-syntactic organization of language systems. By the way, the strong and complicated morpho-syntactic structure of the Georgian verbs is a matter of considerable difficulty. In this connection we offer the following basic suggestions: in Georgian the class of verbs should be divided into two classes – the class of the verbs of simple understanding and the class of the verbs of integrated understanding. At the same time the verbs with integrated understanding are the ones that may be defined on the basis of the verbs with simple understanding. Here, the formal notation theory is used as a tool for the making above mentioned definition by means of which the understanding of the words of integrated understandability is deduced to the understanding of the words of simple understandability. Non-logical-contradictionality of the formal operation of substitution (pertaining to formalism) as well-foundness of the such non-contrasted definitions is relied by notation theory..