

არავერბალური და ვერბალური ხატულების მორფო-სემანტიკური ანალიზი

ნათია ამაღლობელი, ნათია ფუტკარაძე

ილია ჭავჭავაძის სახელმწიფო უნივერსიტეტი, თსუ
natiama2000@yahoo.fr, natia.putkaradze@gmail.com

ხატულა ელექტრონული მეტყველებისთვის დამახასიათებელი ნიშანია, რომელიც ვირტუალურ მიმოწერაში ემოციებისა და გრძობების გამოხატვის საშუალებას იძლევა. ამის გამო მას ხშირად განსაზღვრავენ როგორც პარალინგვისტური მონაცემების წარმოდგენის ხერხს.

დღეისთვის ხატულების სამი ტიპი შეიძლება გამოვყოთ: პუნქტუაციური (“-“), გრაფიკული (☺) და ვერბალური („**გახარებული სმაილიკი**:“) ხატულები.

არსებობს პუნქტუაციური ხატულების ორი ძირითადი სტილი: დასავლური – კლასიკური და ყველაზე პოპულარული (“-“) და აღმოსავლური ანუ **მანგას** სტილი (^_^).

პუნქტუაციური ხატულის სტრუქტურა შეიძლება განვიხილოთ, როგორც სინტაგმატურ ურთიერთდამოკიდებულებაში მყოფი გრაფემებისგან შემდგარი გარკვეული აზრისა და ფუნქციის მქონე ერთეული. შეიძლება გამოვყოთ სამი ძირითადი გრაფემა პირის, თვალებისა და ცხვირის ადგილის გამოსახატავად. აღნიშნული გრაფემები თითქმის იმგვარადვე ფუნქციონირებს, როგორც მორფემები ჩვეულებრივ სიტყვაში. იმისდა მიხედვით, თუ როგორია ამა თუ იმ გრაფემა/მორფემის სემანტიკური დატვირთვა, ხშირია მისი გაორმაგება, აბრევიაცია თუ რედუქცია. დასავლურ და აღმოსავლურ ხატულებში სემანტიკური დატვირთვა ხშირად სხვადასხვა გრაფემა/მორფემაზე მოდის, რასაც საფუძვლად ხალხებს შორის კულტურული განსხვავება უდევს.

გამოიყოფა ხატულების ოთხი ძირითადი ფუნქცია: 1. ექსპრესიული 2. მიმართებითი (იგულისხმება ავტორსა და ადრესატს შორის ურთიერთდამოკიდებულება), 3. თავაზიანობის 4. ირონიისა და იუმორის გამომხატველი.

კლასიკური, არავერბალური იკონური გამოსახულებების პარალელურად ელექტრონულ მეტყველებაში თანდათან აქტიურდება ვერბალური ხატულებიც, რომლებიც ამა თუ იმ ემოციისა და ექსტრის სიტყვიერი გამოსატყობით არის მიღებული; მხედველობაში გვაქვს შემდეგი ტიპის აღწერითი ხატულები: „**გახარებული სმაილიკი**:, **ბედნიერი სმაილიკი**:, მოგყვებიი, მეშველაა, ვთხოვდებიი **პატარძალი სმაილიკი**:“ <http://drive.ge/forum/lofiversion/index.php/t26.html>.

სამეცნიერო ლიტერატურაში გამოთქმულია მოსაზრება, რომ „ემოციათა გამომხატველი ლექსიკა ქართულში (და არა მხოლოდ ქართულში) გარკვეულ ცვლილებებს განიცდის აღწერილობითიდან გამოსახულებითისაკენ. ხდება ექსპრესიის უნიფიცირება, მაგ., „მოწყენა“ ყველამ ერთნაირად მოხაზული სემიოტიკური სიმბოლოთი უნდა წარმოდგინოს; ადამიანი იმდენად ეჩვენა შეტყობინების ვიზუა-

ლიზაციას, რომ ტექსტის აღქმისას ნაკლებად ძაბავს წარმოსახვას, მას ურჩევნია, ემოციაც კი გრძობითიდან ფიზიკურ, მატერიალურ არსამდე დაიყვანოს“ (მ. ტაბიძე); თუმცა, ინტერნეტფორუმების ქართულის (და არა მხოლოდ ქართულის) ანალიზმა ცხადყო, რომ გრაფიკული სიმბოლოების განვითარების შემდეგი ეტაპი სწორედ აღწერითი ხატულებია; როგორც ჩანს, ქართული ელექტრონული დისკურსის მონაწილეებს არ ჰყოფნით სტანდარტული იკონური გამოსახულებები გრძობისა თუ კონკრეტული სიტუაციის გამოსახატავად; გარდა ამისა, მსგავსი „შემოქმედებითობა“ მათ ელექტრონულ გზავნილებს უნიკალურსა და სხვებისგან გამოჩეულს ხდის.

გამოყოფთ სახელური ფრაზის (NP) ტიპისა და ზმნური ფრაზის (VP) ტიპის ვერბალურ ხატულებს. მაგ., NP: **ჩემი გაოგნებული სმაილიკი;** VP: **სერვერი შემომეფორმატას სმაილიკი;** **კიდევ დავაბეჭდებ სმაილიკი;** **ხომმაგარი დავაკომენტარებს სმაილიკი;** და სხვ. ზმნური ფრაზის შემცველი ხატულების ანალიზისას იკვეთება ზმნათა რამდენიმე ყალიბი; ფორუმის მომხმარებლები ქართული შესატყვისის არმქონე ინგლისურ ტერმინებს რამდენიმე კონკრეტულ ზმნურ ყალიბს არგებენ. ირკვევა, რომ ზმნისწინთა არჩევანიც შეზღუდულია... ვხვდებით ორთოგრაფიულ შეცდომებსაც, მაგ., **წინდოუსის გადაყენების სმაილიკი;** სადაც მომხმარებელი ქართული ანბანით წერისას **„წინდოუსის“** ნაცვლად ბეჭდავს **„წინდოუსს“**. მსგავსი შეცდომების სიმრავლე, ვფიქრობთ, კონკრეტული სიტყვების დაწერილობით უნდა აიხსნას; ფორუმის მომხმარებელი, ამ შემთხვევაში, კომპიუტერული პროგრამის ინგლისური სახელწოდების (“Windows”) გაგლენით კლავიატურაზე კრეფს “w” ასოს, რაც ქართული შრიფტით წერისას სწორედ „წ“-ს შეესაბამება, შესაბამისად, „ვ“ ასო-ნიშნის ნაცვლად გვხვდება „წ“.

ოდენ სახელური ფრაზის შემცველი ღიმილაკები პროცენტულად გაცილებით აღემატება ზმნური ფრაზით გამოხატულ აღწერით ხატულებს. საერთოდ კი, იმდენად პროდუქტიულია ელექტრონული მეტყველების ეს „ჰიბრიდული“ მონაცემი, რომ თავისუფლად შეიძლება საკმაოდ ვრცელი კორპუსისა და ლექსიკონის შედგენაც.

Morpho-semantic analysis of non-verbal and verbal emoticons

Natia Amaghlobeli, Natia Putkaradze

Ilia Chavchavadze State University, Tbilisi State University

natiama2000@yahoo.fr natia.putkaradze@gmail.com

Emoticons are virtual symbols which represent feelings or emotions in computer-mediated correspondence. They are often classed as a means of representation of non-verbal paralinguistic cues.

Three types of emoticons can be differentiated: punctuation (“;-)”), graphic (☺) and verbal (“:gaḥarebuli smailiki:” “Happy Smiley”) emoticons.

There are two main styles of punctuation emoticons: Western – classical and the most popular (“;-)”), and Eastern or Manga style (^_^).

The structure of smiley can be considered as a syntagmatic entity of the graphemes with a certain function and meaning. These graphemes function in almost the same way as the morphemes in the word. Three main graphemes are distinguished: for a nose, for an eye and for a mouth. According to their semantic importance graphemes can be redoubled or abbreviated. Due to the cultural differences between Asian and European peoples, in Western and Eastern emoticons they attach different semantic importance to the same graphemes/morphemes.

Four main functions of a smiley can be distinguished: expression of emotion, irony, relationships and politeness.

The verbal emoticons become more and more productive in parallel with the classical, non-verbal emoticons. The verbal emoticons express the emotions and gesticulations using the words; e.g.: “:gaḥarebuli smailiki:, :bednieri smailiki:, mogqvebiii, mešvelaa, vthovdebiii, :paṭarġali smailiki” / “:Pleased smiley:, :happy smiley:, I’ll marry you, I’m saved, :bride smiley:” <http://drive.ge/forum/lofiversion/index.php/t26.html>.

According to linguistic studies, the descriptive character of the words expressing emotions tends to be depicted and expressed by the images in Georgian and other languages... Expression tends to become unified, e.g.: “boredom” should be associated with the particular semiotic sign for everyone. “A human-being gets used to the visualisation of messages so much that while perceiving the text they strain their imagination to a lesser extent. A human being prefers to reduce emotion from sensual to physical, material essence” (M. Tabidze); However, analysis of Georgian internet forums revealed that the next stage of the development of descriptive emoticons is graphical symbols. Obviously, standard iconic

images are not sufficient for the creators of Georgian electronic discourse; in order to express their specific emotions the participants of the forum use words in the format of the smiley by this “creativity” making their electronic messages unique and different from others.

We differentiate two types of verbal emoticons: emoticons containing Noun Phrase (NP) and those containing Verbal phrase (VP); e.g.: **NP:** ”:čemigaognebulismailiki:“ / ”:My astonished Smiley:“; **VP:** ”:serverišemomeformatasmailiki:“ / ”Accidently I’ve formatted Server smiley“, ”:kidevdavabekafebsmailiki:“ / ”I’ll back it up again smiley“, ”:hommagaridavakomentaressmailiki:“ / ”I’ve made a cool comment smiley, haven’t I? And so on

Arguably, several “moulds” of verbs were identified while analysing VP emoticons. Obviously, Georgian participants of forums use several types of verb-formation while seeking and establishing Georgian equivalents to the English terms . Clearly, not all the pre-verbs can be used in the process of the formation of verbal emoticons. In addition, the authors came across a great number of spelling mistakes, such as: ”:čindousisgadaqenebissmailiki:“ – The users of the forum type ”čindous“ instead of ”windows“; can be explained by the spelling of this particular word. The forum user, in this particular case, types ”w“ in order to write ”windows“; but, because of the automatic Georgian Unicode converter, the letter “č” appears on the comment-walls of the Georgian Internet-forums.

The research showed that the percentage of the NP Emoticons outnumbers the VP emoticons. Generally this “Hybrid” of the electronic discourse proves to be exceedingly productive which makes possibility of creating a dictionary based on the body of information very real.

ქართული ენის კომპიუტერული სუფლიორი¹

ნინო ამირეზაშვილი, ლიანა ლორთქიფანიძე, ლიანა სამსონაძე,
ნინო ჯავაშვილი

სსიპ არჩილ ელიაშვილის მართვის სისტემების ინსტიტუტი
ninomaskh@yahoo.com, L_lordkipanidze@yahoo.com, liasams@yahoo.com,
ninojavashvili@yahoo.com

ქართული ენის კომპიუტერული სუფლიორი უნარდაქვეითებული პირებისთვის ქართულად წერის ხელშემწყობი კომპიუტერული პროგრამაა, რომელიც, რამდენადაც ჩვენთვის ცნობილია, საქართველოში ჯერ არ დამუშავებულა. პროგრამა მომხმარებელს სიტყვის სწორ ვარიანტს სთავაზობს და კლავიატურის მინიმალური გამოყენებით უადვილებს კომპიუტერთან ურთიერთობას. შედეგის სამეფო ტექნიკური უნივერსიტეტის (Dept. of Speech, Music and Hearing at the Royal Institute of Technology (KTH) in Sweden) მეცნიერთა ჯგუფი ასეთი სისტემის შექმნაზე წლების განმავლობაში მუშაობდა. ეს სისტემა **Prophet**-ის სახელითაა ცნობილი. **Prophet**-ი პროგრამაა, რომელიც კომპიუტერთან მუშაობის დროს მომხმარებელს სიტყვის სწორ ვარიანტს სთავაზობს წინამორბედი სიტყვიდან ან ასოდან გამომდინარე. გადაწყვეტიტ ჩვენც შეგვექმნა **Prophet**-ის ანალოგიური ქართული სისტემა.

Prophet-ის ქართული ვერსიისათვის, რომელსაც ჩვენ ვუწოდებთ “ქართული ენის კომპიუტერულ სუფლიორს”, საჭიროა: **Prophet**-ის პროგრამული კოდის ქართულ ენაზე მორგება; ქართული ტექსტური კორპუსის შედგენა (არანაკლებ მილიონი სიტყვისა); **Prophet**-ის ქართული სალექსიკონო ბაზების შევსება; ქართული ენის მორფოლოგიური პროცესორის შემუშავება; *ავიქსების* ლექსიკონის შედგენა და **Prophet**-ის სალექსიკონო ბაზების მოდიფიკაცია.

ამ დროისთვის ქართული ტექსტური კორპუსი შევსებულია დაახლოებით მილიონი სიტყვით. კორპუსის შესადგენად გამოვიყენეთ ქართული ინტერნეტგვერდები. მასალა ერთგვაროვანი რომ არ ყოფილიყო, მოვიძიეთ სხვადასხვა თემატიკის ტექსტები თანაბარი რაოდენობით. დავამუშავეთ თორმეტი თემატიკა, ესენია: საქართველოს ისტორია, რელიგია, კულტურა, მედიცინა, სპორტი, ეკონომიკა, პოლიტიკა, შოუბიზნესი, ოჯახი და ბავშვები, ხალხი და საზოგადოება. ტექსტური კორპუსი შემდგომში სათანადო დამუშავებას მოითხოვს. შესაყვანი ფაილის ფორმატი უნდა იყოს ერთი სიტყვა თითოეულ სტრიქონზე. სამწუხაროდ, ონლაინ ტექსტებში ძალზე ხშირია სხვადასხვა სახის შეცდომები (სიტყვის შეცდომით ჩაწერა, არა-

¹ პროექტი ხორციელდება სსიპ – ქართველოლოგიის, ჰუმანიტარული და სოციალური მეცნიერებების ფონდში (რუსთაველის ფონდი) მოპოვებული გრანტის მეშვეობით (გრანტი № 064-08-1-030). წინამდებარე პუბლიკაციაში გამოთქმული ნებისმიერი აზრი ეკუთვნის ავტორს და შესაძლოა არ ასახავდეს სსიპ - ქართველოლოგიის, ჰუმანიტარული და სოციალური მეცნიერებების ფონდის (რუსთაველის ფონდი) შეხედულებებს.

ფორმატირებული ტექსტი განსხვავებული ზედმეტი გამოტოვებული ადგილებით, ზედმეტი პუნქტუაციური ნიშნები და ა.შ.). **Prophet**-ისთვის სათანადო სახის მისაცემად კორპუსზე ჩავატარეთ სხვადასხვა სახის დამატებითი სამუშაოებიც.

Prophet-ის პროგრამული უზრუნველყოფის გამოყენებით ტექსტური კორპუსიდან უკვე შევსებულია ქართული სალექსიკონო ბაზები. ესენია: 1) **Pair** (წყვილების *ლექსიკონი*; 2) **First** *ლექსიკონი* და 3) **Main** *ლექსიკონი*.

1) **Pair** *ლექსიკონი* შედგება ქართული კორპუსიდან ამოკრებილი იმ სიტყვათა წყვილებისაგან, რომლებიც ტექსტში ერთმანეთის მიყოლებით გვხვდება. მათ მიწერილი აქვთ კორპუსში შეხვედრის სიხშირე.

2) **Main** *ლექსიკონი* შედგება ტექსტური კორპუსიდან ამოკრებილი სიტყვებისაგან, რომლებსაც, აგრეთვე, მიწერილი აქვთ შესაბამისი სიხშირე.

3) **First** *ლექსიკონი* შედგება ანბანის თითოეული ასოს მიხედვით ამორჩეული ისეთი 9 სიტყვისაგან, რომელიც ყველაზე ხშირად გვხვდება ტექსტურ კორპუსში. მათი კომპილირება შეიძლება **Main** *ლექსიკონი*დან.

ამავე დროს იქმნება ენის მორფოლოგიური პროცესორი, რომელიც შემდეგ ეტაპზე იძლევა: 1. **ავიქსების** *ლექსიკონის* შედგენისა და 2. **Main** *ლექსიკონის* მოდიფიკაციის საშუალებას.

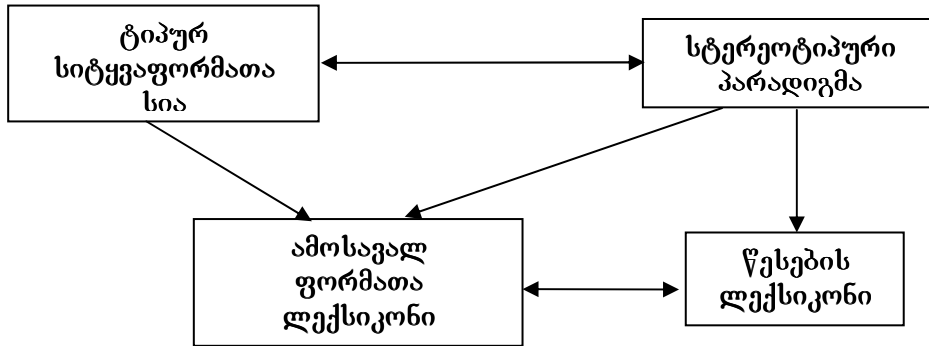
1. **ავიქსების** *ლექსიკონში* ჩაიწერება **Main** *ლექსიკონში* არსებული ყველა სიტყვის შესაბამისი პარადიგმის მაწარმოებელი წესის კოდი და ავიქსების თანამიმდევრობა.

2. **Main** *ლექსიკონის* მოდიფიკაცია გულისხმობს *ლექსიკონის* თითოეული სიტყვისათვის შესაბამისი პარადიგმის მაწარმოებელი წესის კოდის მიწერას.

ქართული მორფოლოგიური პროცესორის შემუშავება ხორციელდება ცოდნის დაგროვების **MESLM** მულტიენობრივი ექსპერტული სისტემის საშუალებით, რომელიც წარმატებით იქნა გამოყენებული **Prophet**-ის რუსული ვერსიისთვის. **MESLM** სისტემა შეიძლება განხილულ იქნეს როგორც ინსტრუმენტი, რომელიც ბაძავს ადამიანის აზროვნებას და ქმნის მორფოლოგიურ წესებს ნებისმიერი ენის პროცესორისათვის. მართალია, **AI** მეთოდები ბუნებრივი ენების დასამუშავებლად ადრეც გამოიყენებოდა, მაგრამ მათი საშუალებით მორფოლოგიური პროცესორის შემუშავება საკმაოდ რთული იყო, განსაკუთრებით, აგლუტინაციური და ფლექსიური ენებისათვის. ასეთ სისტემებში ექსპერტების ცოდნის გამოყენება მხოლოდ სპეციალური ოპერატორების ან ხელოვნური ენების ათვისებითაა შესაძლებელი, რაც სრულიად ზედმეტია **MESLM** სისტემის მოხმარებისას. აქ ცოდნის შექმნა შესაძლებელია ენის ამოსავალი სიტყვების სტერეოტიპური პარადიგმების საფუძველზე. პარადიგმის ტექსტის სეგმენტირება სიტყვებად და არალექსიკურ მასალად სახეთა ამოცნობის მეთოდების გამოყენებით ხდება. შემდეგ განისაზღვრება ლემა და პარადიგმის თითოეული სიტყვაფორმისთვის შეიქმნება მორფოლოგიური წესი.

MESLM სისტემის პროგრამული უზრუნველყოფის დახმარებით ქართული ენის მორფოლოგიური პროცესორის შემუშავება მოიცავს სამ ძირითად კომპონენტს. ესენია: ზმნა, არსებითი სახელი და დანარჩენი მეტყველების ნაწილები. სა-

მივე შემთხვევაში ექსპერტული სისტემის მიერ ცოდნის შექმნა ერთი და იმავე სქემით ხორციელდება (იხ. ნახ. 1).



ნახ. 1. MESLM ექსპერტული სისტემის მიერ ცოდნის შექმნის სქემა

აღნიშნული პროცესი განვიხილოთ ზმნის მაგალითზე. ცოდნის შექმნის მიზნით სისტემას უნდა მიეწოდოს ზმნის **სტერეოტიპური პარადიგმა**. დავუშვათ, ეს ზმნაა – **‘ენანება’**. სისტემა **სტერეოტიპურ პარადიგმაზე** დაყრდნობით შეიძენს ცოდნას შესაბამისი მორფოლოგიური წესის შესახებ; შეინახავს მას **ამოსავალ ფორმათა ლექსიკონში** და **წესების ლექსიკონში**. იმავედროულად ამოსავალ ფორმას მიეწერება სათანადო წესის ნომერი; სისტემა ერთხელ უკვე დამახსოვრებული წესის გამოყენებით იმავე ტიპის სხვა ამოსავალი სიტყვების პარადიგმების შემოწმების საშუალებას იძლევა. მას შემდეგ, რაც ექსპერტი დარწმუნდება, რომ **ტიპურ სიტყვაფორმათა სია** სწორადაა შერჩეული, **ამოსავალ ფორმათა ლექსიკონს MESLM** სისტემის პროგრამული უზრუნველყოფის ბრძანებით დაემატება **ტიპურ სიტყვათა სია**, რომელშიც თითოეულ მათგანს მიწერილი ექნება სათანადო წესის ნომერი:

- ელანდება v1
- ემგვანება v1
- ენანება v1

პროცედურა გამეორდება ყველა მეტყველების ნაწილის **სტერეოტიპური პარადიგმისათვის**. უნდა აღინიშნოს, რომ ლინგვისტს ნებისმიერ დროს შეუძლია შესამოწმებლად გამოიყენოს ერთხელ უკვე ჩაწერილი წესი და შესაბამისად გაზარდოს **ტიპურ სიტყვაფორმათა სია**.

პროექტის განხორციელების შედეგად შეიქმნება **Prophet**-ის ქართული ვერსია. კომპიუტერული პროგრამა დაეხმარება მომხმარებელს ქართული ტექსტების აკრეფის დროს სიტყვათა თანმიმდევრობის მითითებით ამოირჩიოს გრამატიკულად კორექტული სიტყვები. სალექსიკონო ბაზაში შევა 100000 ამოსავალი სიტყვა და მათი ფორმაწარმოების ყველა წესი. ჯერჯერობით ამოსავალ ფორმათა და წესების ლექსიკონი შეესებულება 30000 ერთეულით.

Georgian Computer Prompter¹

N. Amirezashvili, L. Lortkipanidze, L. Samsonadze, N. Javashvili

LEPL Archil Eliashvili Institute of Control Systems

ninomaskh@yahoo.com, L_lordkipanidze@yahoo.com, liasams@yahoo.com,

ninojavashvili@yahoo.com

Georgian computer prompter is a software which can assist the disabled to write in Georgian on the computer. As is known this problem cannot yet be solved by any current software. This system suggests the correct forms of the word and makes it easy to use the keyboard with the minimum of effort. The research group at KTH (Dept. of Speech, Music and Hearing at the Royal Institute of Technology in Sweden) has been working on the assistance system for selection of the desired and correct forms for a long time. Programs that carry out this function are called word predictor "Prophet". A word predictor suggests words whilst a person is writing, either based on the preceding word or the first letter(s) of the current word. We have decided to create a Georgian system similar to Prophet.

For the Georgian version of Prophet which we call "Computer Prompter", it is necessary to adjust the programme code of Prophet to the Georgian language; to create the Georgian text corpora (no less than a million words); to increase the database of the Georgian dictionary; to develop a morphological processor of the Georgian language and to create the dictionary of affixes and modify the dictionary database of Prophet.

By this time the Georgian text corpora has been filled with about one million words. In order to create the text corpora we used Georgian internet sites. In addition, to ensure the variety of the themes, we collected and developed texts comprising twelve themes: Georgian history, Religion, Culture, Medicine, Sport, Economic, Politics, Show-business, Family, Children, People and Society. The text corpora require special processing such as the format of the entry file must have only one word in each line. Unfortunately in online texts there are a lot of mistakes (such as errors in spelling, unformatted texts with different gaps, overload with punctuation symbols, etc). In order to give Prophet the appropriate shape we carried out a different kind of additional work.

¹ The designated project has been made possible by financial support from the Foundation For Georgian Studies, Humanities and Social Sciences (Rustaveli Foundation) (Grant № 064-08-1-030). Any ideas in this publication are those of the authors and may not represent the opinion of Foundation For Georgian Studies, Humanities and Social Sciences (Rustaveli Foundation) itself.

The database of the Georgian dictionary is already filled up from the text corpora by using the Prophet software. These are: 1) The Pairs dictionary; 2) The First dictionary; 3) The Main dictionary.

1. The Pairs dictionary consists of the word pairs selected from the text corpora and used in the text in succession. The frequency of their occurrence in the text is also computed and shown.
2. The Main dictionary consists of all of the words selected from the text corpora followed by the computed relevant frequency.
3. The First dictionary consists of 9 words for each letter of the alphabet, which occur most frequently in the text corpora. It is possible to compile them from the Main dictionary.

At the same time, a Georgian morphological processor is being developed which will make it possible, at the next stage of the project, to create the Dictionary of Affixes and modify the Main Dictionary.

1. In the dictionary of Affixes the code for each rule of derivation of the paradigm and the order of the affixes corresponding to each word included in the Main dictionary will be supplied.
2. Modification of the Main dictionary implies ascription of the code of a rule for derivation of the respective word paradigm.

The development of the Georgian morphological processor is provided by means of knowledge collection MESLM Multilanguage expert system, which was successfully used for the Russian version of Prophet.

MESLM can be regarded as an instrument which imitates the human mind and creates morphological rules for any language processor. As is known, when AI techniques were applied to processing natural languages earlier the development of morphological processes turned out to be quite difficult, especially for agglutinative and flexible languages. It is possible to apply the experts' knowledge to such systems only through adoption of new operators or languages which are not required for the use of the MESLM system. It is possible to acquire knowledge with the help of such systems on the basis of typical paradigms of the key words of the language. Segmentation of the text paradigms into words and non-lexical material is provided through recognition of the images. Next, the lemma is developed and a separate morphological rule for each word form of the paradigm will be developed.

Development of the Georgian language processor by means of the MESLM system software includes three components. These are: verb, noun and other parts of speech. In all the

three cases of acquisition of knowledge by the expert system is provided in accordance with one and the same scheme (See Fig. 1)

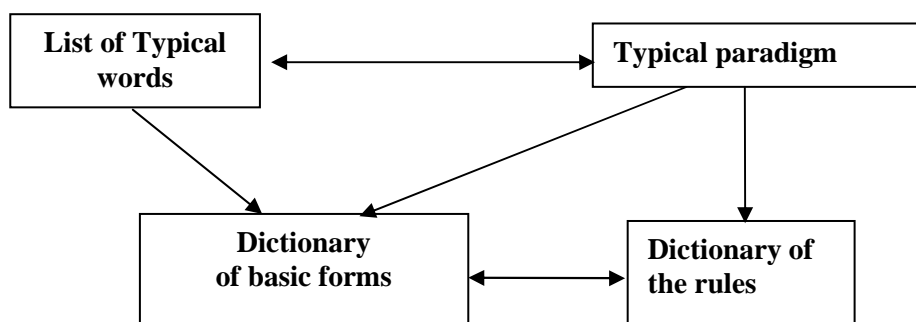


Fig. 1. Acquisition of knowledge by the MESLM expert system

We will illustrate the process discussed above using the example of the verb. With the purpose of acquisition of knowledge the system should be supplied by the typical paradigm of the verb. Assume that this is the verb *enaneba* 'to regret'. Based on the typical paradigm, the system will acquire the knowledge about the relevant morphological rules and next, save it in the Dictionary of basic forms as well as in the Dictionary of the Rules. Simultaneously, the number of the relevant rule will be given to the basic form. The system, using the once saved rule, allows for examination of the paradigms of other basic words of the same type. When the expert makes sure that the list of typical word forms is selected correctly, the List of Typical words will be added to the dictionary of basic words, in the order of MESLM system software, in which, the number of relevant rule will be given to each of them.

elandeba v1 'it seems to him'
emgvaneba v1 'he likens to smb.'
enaneba v1 'he is sorry for smth.'

This procedure will be repeated for the stereotypical paradigm of each of the parts of speech. It should be noted that the linguist will be able to use the rule recorded at any time and accordingly increase the list of typical word forms.

As a result of the Project implementation the Georgian version of Prophet will be created. The software will assist the users to select grammatically correct words while keying in Georgian texts by means of specifying the sequence of words. The database of the dictionary will contain 100,000 basic words and all of the rules for their derivation. At present 30,000 units have been added to the dictionary of basic forms and rules.

ელექტრონული ინგლისურ-ქართულ-რუსულ-გერმანულ-ფრანგულ-იტალიური მათემატიკური ტერმინების სიტყვანი

ჯ. ანთიძე, ც. გაბესკირია, გ. ჯაიანი, გ. დათაშვილი, ი. ნადირაძე

ი. ვეკუას სახ. გამოყენებითი მათემატიკის ინსტიტუტი, თსუ
jeantidze@yahoo.com

მოსვენებაში წარმოდგენილი იქნება მათემატიკური ტერმინების ელექტრონული მრავალენოვანი სიტყვანი. მას აქვს მოხერხებული ინტერფეისი, რომელიც საშუალებას გვაძლევს სწრაფად ვიპოვოთ ამა თუ იმ ტერმინის შესატყვისი მეორე ენაზე, აქვს თითოეული ენის შესატყვისი კლავიატურის მოდელი, რომელზედაც გადართვა ხდება ავტომატურად. არსებობს ახალი სიტყვების დამატების შესაძლებლობა.

სიტყვანი შეიცავს 2000-ზე მეტ ტერმინს და დაფუძნებულია გამოქვეყნებულ სიტყვანზე: Ts. Gabeskiria, G. Jaiani, J. Antidze, G. Datashvili. English-Georgian-Russian-German-French-Italian Glossary of Mathematical Terms, Lecture notes of TICMI, volume 6, 2005, Tbilisi University Press. გათვალისწინებულია სიტყვანის გამოყენების დემონსტრირება.

English-Georgian-Russian-German-French-Italian Electronic Glossary of Mathematical Terms

J. Antidze, Ts. Gabeskiria, G. Datashvili, G. Jaiani and I. Nadiradze

I.Vekua Institute of Applied Mathematics, Tbilisi State University
jeantidze@yahoo.com

This paper will present a new, electronic, multilingual, glossary of mathematical terms. It has a comfortable interface which enables fast searching of a mathematical term requested to be translated from one language into another. The Glossary possesses a modeled keyboard for each language processed within the Glossary. Switching from one language to another is very simple as the corresponding keyboard appears automatically. The Glossary presents a possibility to add new terms as needed .

The Glossary presented in this paper contains more than 2000 terms and is based on the earlier published work of Ts. Gabeskiria, G. Jaiani, J. Antidze and G. Datashvili “English-Georgian-Russian-German-French-Italian Glossary of Mathematical Terms”; Lecture notes of TICMI, volume 6, 2005, Tbilisi University Press.

During the presentation the procedure of usage of the Glossary will also be demonstrated.

ქართული ტექსტების სრული კომპიუტერული მორფოლოგიური და სინტაქსური ანალიზის შესახებ

ჯემალ ანთიძე, ნანა გულუა, დ. მიშელაშვილი, ლ. ნუკრაძე
ივ. ჯავახიშვილის სახელობის თბილისის სახელმწიფო უნივერსიტეტი
jeantidze@yahoo.com, ngulua7@mail.ru

ქართული ტექსტების სრული კომპიუტერული მორფოლოგიური და სინტაქსური ანალიზი წარმოადგენს ერთ-ერთ მთავარ კომპონენტს ისეთი პრობლემების გადასაწყვეტად, როგორცაა ქართული ენიდან მანქანური თარგმანი და ქართული ტექსტების ორთოგრაფიული წესების ავტომატური შემოწმება, ასევე, ხელოვნური ინტელექტის ისეთი პრობლემები, რომლებიც მოითხოვენ ქართული ტექსტების გამოყენებას. ქართული სიტყვების კომპიუტერული სრული მორფოლოგიური ანალიზის სისტემა ჯერჯერობით არ არსებობს. თუ ჩვენ გვსურს ქართული ენა გამოვიყენოთ კომპიუტერთან კომუნიკაციისათვის, ამ პრობლემის გადაწყვეტა ძალზე საშური საქმეა. ამ პრობლემის გადასაწყვეტად სასრული ავტომატების გამოყენება, რაც საყოველთაოდ მიღებულია დასავლეთ ევროპის ენებისათვის, შეუძლებელია იმიტომ, რომ ზოგიერთი ზმნის ფორმა მოითხოვს უკან დაბრუნებას, რაც სასრული ავტომატებით ვერ განხორციელდება. მეორე მხრივ, სრული გადარჩევის ალგორითმის გამოყენება ანელებს სიტყვის მორფემებად დაშლის პროცესს. ამიტომ ჩვენ გამოვიყენეთ მეთოდი, რომელიც აჩქარებს ძებნის პროცესს სრული გადარჩევის ალგორითმის გამოყენებასთან შედარებით. ეს მეთოდი იყენებს შეზღუდვებს სწორი მორფემების ამოსარჩევად. მოცემული მომენტისათვის მოწმდება სიტყვიდან უკვე გამოყოფილი მორფემები აკმაყოფილებს თუ არა მოცემულ შეზღუდვას. თუ შეზღუდვა სრულდება, მაშინ ძებნის პროცესი გრძელდება, წინააღმდეგ შემთხვევაში ბოლო გამოყოფილი მორფემის უკუგდება ხდება და უკან ვბრუნდებით ახალი ალტერნატივის ასარჩევად. ასეთი გზით არასწორად არჩეული მორფემების უკუგდება ხდება ადრეულ ეტაპზე, რაც აჩქარებს ძებნის პროცესს.

შეზღუდვები არის ლოგიკური გამოსახულებები, რომლებიც შეიძლება შედგენილ იქნეს მორფემათა თვისებებზე ქართული ენის მორფოლოგიის წესების დაცვით. სრულ კომპიუტერულ მორფოლოგიურ ანალიზში ჩვენ ვგულისხმობთ მოცემული სიტყვის ყველა შესაძლო სწორ დაშლას მორფემებად და ყოველი სწორი დაშლისათვის მორფოლოგიური კატეგორიების დადგენას. ასეთი განსაზღვრება შეიცავს სიტყვების ომონიმას. ქართულ ენაში ფართოდაა გავრცელებული შემდეგი სახის ომონიმები:

1. სხვადასხვა ზმნის ფორმების დამთხვევა, როგორცაა აწმყოს წრის ფორმების დამთხვევა. მაგალითად, “აგებს” შეიძლება აღებული იყოს “აგება” მნიშვნელობით ან “წაგება” მნიშვნელობით და ა.შ..

2. ზმნის ფორმის დამთხვევა საწყისის ფორმასთან. მაგალითად, “ამოხსნა” შეიძლება იყოს საწყისის ფორმა ან ზმნის ფორმა - “მან ამოხსნა”.

3. სიტყვაში სხვადასხვა კლასის მეზობელი მორფემების დამთხვევა. მაგალითად, “ა” როგორც ზმნისწინი ან ხმოვანი პრეფიქსი ანდა ძირის პირველი ასო ზმნის შემდეგ ფორმებში: “ა-ა-ალებს”, “ა-ალებს” და “აღდება”. როცა ჩვენ ვხედავთ “ააა-ლებს” ზმნის ფორმის პირველ ასოს, ჩვენ არ შეგვიძლია დავასკვნათ, რომელი მორფემაა, სანამ არ დავინახავთ შემდეგ ორ ასოს. პირველ სიტყვაში პირველი “ა” არის ზმნისწინი, მეორეში – ხმოვანი პრეფიქსი და მესამეში – ძირის პირველი ასო. ეს ნიშნავს, რომ ქართული სიტყვების მორფემებად დახლეჩა საჭიროებს ყველაზე მცირე გარჩევის ალგორითმს LL(2) გრამატიკისათვის. ე.ი. ქართული სიტყვების სრული მორფოლოგიური ანალიზი სასრული ავტომატით შეუძლებელია.

მეორე შემთხვევაში მორფოლოგიურმა ანალიზმა ზმნის ფორმისათვის “ამოხსნა” უნდა მოგვცეს ორი განსხვავებული გარჩევა, ერთი – ზმნისათვის და მეორე – საწყისისათვის. ამისათვის გვჭირდება არადეტერმინისტული ალგორითმი. დეტერმინისტული ალგორითმი ვერ მოგვცემს ორ განსხვავებულ გარჩევას ერთი და იმავე სიტყვისათვის. ამგვარად, დეტერმინისტული ალგორითმი არ გამოდგება. მრავალ ავტორს ქართული სიტყვების კომპიუტერული მორფოლოგიური ანალიზი შესრულებული აქვს ან სასრული ავტომატით ანდა არადეტერმინისტული ალგორითმით. სრული მორფოლოგიური ანალიზისათვის ჩვენ უნდა გამოვიყენოთ არადეტერმინისტული ალგორითმი, მაგალითად, მარცხნიდან მარჯვნივ სიღრმეში ძებნის ალგორითმი უკან დაბრუნებებით. რადგან უკან დაბრუნებები ანელევენ ალგორითმის სისწრაფეს, ჩვენ უნდა ვიპოვოთ მეთოდი, რომელიც შეამცირებს უკან დაბრუნების შემთხვევებს. ასეთი საშუალება არსებობს. ჩვენ შეგვიძლია გამოვიცხოთ ის მორფემები, რომლებიც კონფლიქტში მოდიან მოცემულ მომენტში ნაპოვნ მორფემებთან. მეორე მხრივ, ჩვენ შეგვიძლია დავყოთ მორფემები კლასებად ისე, რომ თითოეული კლასის წარმომადგენელი ერთხელ შეგვხვდეს სიტყვაში. მორფემათა კლასებს შორის მნიშვნელოვანია ძირები. ძირები შეიძლება დაიყოს კლასებად ისე, რომ ყოველი მორფემა, რომელიც შეგვხვდება ზმნის ფორმაში, ცალსახად განსაზღვრავდეს ზმნის ფორმის მორფოლოგიურ კატეგორიას. ყველაფერი ეს ამცირებს უკან დაბრუნების შემთხვევებს. ამის შემდეგ სიტყვის მორფოლოგიური კატეგორიების დადგენა ადვილია. ჩვენ განვახორციელებთ სრული მორფოლოგიური ანალიზი ჩვენი პროგრამული საშუალებით და დ. მელიქიშვილის მიერ წარმოდგენილი ზმნების კლასიფიკაციის გამოყენებით. ექსპერიმენტული სინტაქსური ანალიზი განვახორციელებთ ჩვენ მიერ შექმნილი სინტაქსური ანალიზის პროგრამული საშუალებით. იგი აადვილებს ქართული წინადადებების სინტაქსური წესების აღწერას.

Towards the Complete, Computer Morphological and Syntactic Analysis of Georgian Texts

J.Antidze, N.Gulua, D.Mishelashvili, L.Nukradze

Tbilisi State University

jeantidze@yahoo.com, ngulua7@mail.ru

The complete, computer morphological and syntactic analysis of Georgian texts is one of the main components provided to solve such problems as machine translation from Georgian into another language. In addition, those problems of the automated spell check of Georgian texts, as well as the problems related to artificial intelligence, which require computer processing of Georgian texts. The system which would provide complete computer morphological analysis of Georgian words has not yet been developed. In order to use the Georgian language for the purpose of computer-mediated communication, the above mentioned problem needs an urgent solution.

Finite automation which is widely used for Western European languages, does not prove to be feasible in the case of Georgian. This is because some verb-forms of the Georgian language require back-tracking which cannot be achieved by finite automation. On the other hand the application of a full search algorithm slows the process of morphological analysis. For these reasons we created a method which speeds up the analysis compared to that of a full search algorithm. This method applies constraints while selecting the correct morpheme. Having been separated from the word the morphemes are checked regarding the satisfaction of their constraints by means of morphological analysis. If the constraint is satisfied, the tool continues the process of separation of other morphemes. In the opposite direction it performs backtracking in order to search possible new alternatives and thus rejects the last separated morpheme. In this way, the incorrect alternatives are removed at an initial stage which speeds up the process of search.

The constraints are logical expressions, which can be based on the features of morphemes composed according to the rules regulating Georgian morphology. By the term "complete computer morphological analysis" we recognise all valid variants arising while splitting the word-form into morphemes as well as the establishment of morphological categories for each of the cases of splitting. Such definition contains word homonymy. The following types of homonyms are widespread in Georgian:

1. Graphic coincidence of different verb-forms such as coincidence in verbs of the Present circle. For instance, verb-form "agebs", which may denote the following concepts: "lose" (for instance, money) or "build" (for instance, a plan)

-
2. Graphical coincidence of a verb-form with its infinitive. For instance, "amoxsna" may mean "resolution" or "he has resolved";
 3. While splitting the verb-form, graphic coincidence of morphemes from different cluster classes, for instance, "a" is used either as the preverbal or vowel prefix or as the first letter of a verb's root in the following verb-forms: "a-a-alebs", "a-alebs" and "aldeba". Having seen the first letter of the verb-form "aaalebs", we cannot state which morpheme is actualised in the context until we have seen the following two letters. In the first example, the first "a" is used as a preverbal. In the second example, the first "a" is a vowel prefix and in the third example, the first "a" is the first letter of the verb-root "al". This means that in order to split Georgian verbs into morphemes at least parsing algorithm for LL(2) grammar is necessary i.e. complete morphological analysis of Georgian words by finite automaton is definitely impossible.

In the second case, a morphological analysis of the verb-form "amoxsna" must present two different parsings: one for the infinitive and the second for the verb-form. For this purpose a non-deterministic algorithm is necessary as the deterministic algorithm cannot give two different parses for one and the same word-form. Therefore, a deterministic algorithm is not valid for complete morphological analysis of Georgian words. Many authors have performed the morphological analyses of Georgian words relying on finite automation or the deterministic algorithm. For a complete morphological analysis, the non-deterministic algorithm should be applied, for instance, such as from left to right, in-depth search algorithm with back-trackings. As the back-trackings are known to reduce the speed of the algorithm, a method which would reduce their number should be found for the analysis. Such a possibility exists. We can exclude those morphemes which come into conflict with the morphemes found at a definite moment. In other cases, we can divide morphemes in classes so that one representative of each class is met in a word-form only once. Roots are considered to be of considerable importance among the word forms. We can divide roots into classes so that each morpheme in a word-form will indicate a definite morphological category. This will arguably reduce back-trackings and make it easy to establish morphological categories. We performed the complete morphological analysis of Georgian words relying on the program created by us and following the classification of verbs determined by D.Melikishvili.

In addition we also performed experimental syntactic analysis of Georgian phrases based on the software program created by us for syntactic analysis which simplifies the description of syntactic rules in Georgian.

**კავკასიური ენების მონაცემთა ბაზა
(საცდელი პროექტი)**

კარინა ვამლინგი, იოჰან დალი, მანანა კობაიძე
karina.vamling@mah.se
manana.kock.kobaidze@mah.se

კავკასიური ენების მონაცემთა ბაზა, რომელიც რამდენიმე წლის წინ შეიქმნა ლუნდის უნივერსიტეტის ენათმეცნიერების ინსტიტუტში, იყო რამდენიმე კავკასიური ენის მორფოლოგიური კატეგორიების აღწერის ცდა. ბაზის მომხმარებლად ივარაუდებოდა ენათმეცნიერები და აგრეთვე, ენის ზოგადი საკითხებით დაინტერესებული არასპეციალისტები.

ბაზა შეიქმნა მცირე მოცულობის მასალაზე. სულ დაახლოებით ექვსასი წინადადებაა წარმოდგენილი, კერძოდ, ქართულის, მეგრულის, ყაბარდოულის და ადიღურ ენებზე. მასალა აღებულია გამოქვეყნებული სტატიებიდან.

მიზანი იყო, რომ შექმნილიყო ამ ენებში არსებული მორფოლოგიური კატეგორიების მომცველი გლოსირების სისტემა. გლოსირება მთლიანად ხელით განხორციელდა.

მასალა წარმოდგენილია მხოლოდ ლათინური ტრანსლიტერაციით. ახლავს ინგლისური თარგმანი. ძიება შესაძლებელია გრამატიკული კატეგორიების მიხედვით, აგრეთვე აფიქსების მიხედვით, ფუძეების მიხედვით და თარგმანში წარმოდგენილი ფუძეების მიხედვით.

მასალის მორფემული თარგმანის შექმნა გაცილებით მარტივი იყო ყაბარდოულისათვის, სადაც თითოეული გრამატიკული კატეგორია თავისი შესაბამისი ნიშნით არის წარმოდგენილი. ქართულსა და მეგრულში, სადაც ასეთი შესაბამისობა ხშირად არ არსებობს, საჭირო გახდა ორმაგი გლოსირება. ერთი მხრივ, თითოეული მორფემის თარგმანი, მეორე მხრივ კი, იმ გრამატიკული კატეგორიების აღნიშვნა, რომლებიც მთლიანად სიტყვაფორმით არის გამოხატული. მაგალითად, ზმნური ფორმა *შეჭამა* დაშლილია ასე:

შე	ჭამ	ა
ზმნისწინი	ძირი	III სუბ. პირის ნიშანი
S3მხ.O3 მხ. ჭამა.წყვეტილი		

ასეთი ორმაგი ანალიზი მომხმარებელს საშუალებას აძლევს ეძიოს მასალა როგორც მორფემების, ასევე გრამატიკული კატეგორიების მიხედვითაც.

ზოგ შემთხვევაში საჭირო გახდა მითითებული ყოფილიყო მორფემის თავდაპირველი ფუნქცია, მაგალითად, ხმოვანი პრეფიქსი მოქმედებითი გვარის პირველი თემობითის ფორმებში აღნიშნული იქნა, როგორც წარმოშობით ქცევის ნიშანი. ზოგ შემთხვევაში საჭირო გახდა ფუნქციაზე მითითება, მაგალითად, პირველი და მეორე პირის ნაცვალსახელები ზმნის არგუმენტების როლში ფორმაუცვლელი სიტყვები არიან. იმისათვის, რომ მომხმარებელს არ გამოჩენოდა მათი ფუნქცია, ამ

ნაცვალსახელებთანაც მიეთითა, როდის გამოხატავენ ისინი სახელობითის, მოთხრობითის თუ მიცემითის ფუნქციას.

მ-სერიის პირის ნიშნები ყველგან მარკირებული იქნა, როგორც ობიექტური პირის ნიშნები, ვ-სერიისა კი – როგორც სუბიექტური პირის ნიშნები. ამავე დროს, ინვერსიული ზმნები აღინიშნა ცალკე ნიშნით.

ის დაახლოებით 150 კატეგორია, რომელიც ამ მონაცემთა ბაზაში არის გამოყენებული, დაჯგუფებულია იმის მიხედვითაც, არის თუ არა იგი ზმნური ან სახელური კატეგორია, ანდა სიტყვაწარმოების თუ ფორმაწარმოების აფიქსი.

Database for Caucasian Languages (Piloting Project)

Karina Vamling, Johan Dahl, Manana Kobaidze

karina.vamling@mah.se

manana.kock.kobaidze@mah.se

Our presentation concerns the database that has been developed at the Department of Linguistics at Lund University by Karina Vamling, Manana Kock Kobaidze, and Johan Dahl (research engineer) in 2002. The project received funding from Elisabeth Rausings minnesfond (Elisabeth Rausing Memorial Fund). Project leader: Karina Vamling.

The aim of the database is to make available on the internet fully glossed examples in several Caucasian languages, showing a number of different grammatical constructions such as morphological causatives, ergative clauses, inversive constructions, evidentials, complex cross-reference marking. The database includes over 600 sentences in Old and Modern Georgian, as well as Megrelian, Kabardian and Adyghe.

The material is presented only in transliteration of the original scripts. The material is provided both with morpheme translation and translation into English.

The project aims neither at automatic analysis nor inclusion of large texts in the database. The aim is to test the possibility of creating a unified annotation for different Caucasian languages.

The linguistic material in the database has been taken from published articles and books. Annotation has been done manually.

The examples are accessible by search for

- (1) grammatical categories,
- (2) prefixes and suffixes,
- (3) roots of the original language,
- (4) stems of the language of translation (English).

Special attention has been given to unified principles of glossing. A general feature of Kartvelian verbal morphology that poses an interesting problem and called for a special solution is that morpheme segmentation is often not sufficient in order to show all the grammatical categories present in the forms. In the database, this has been solved by introducing double representations: both glosses with full information about categories as well as morpheme segmentation. This is illustrated in the following Old Georgian example:

mgel-man	shech'ama	(she-ch'am-a)	sxovar-i
wolf-ERG	S3SG.O3SG.eat.AOR	PRV-eat-S3SG	lamb-NOM
'The wolf ate the lamb'			

The double annotation makes it possible to search not only for morphological categories that are presented by morphological markers, but also for syntactic categories that can be left without morphological marking.

In certain cases the analysis given in the database does not in all details coincide with the analysis presented in the published materials.

- Some abbreviations have been changed in order to unify the name of the category that was presented with different names in different articles.
- Some abbreviations have been changed in order to differentiate the name of different categories that were presented with one and the same label in different articles.
- In some cases, zero markers (e. g. present tense marker in Kabardian, the nominative case marker after stems ending in vowels in Georgian etc.) have been added to the analyzed forms.
- The morpheme analysis of Georgian verbs and, partly, of nominals, presented in database, is made particularly for this database.
- Certain morphemes that have not retained their original function, but are still presented in verbs have been marked separately. For instance, the vowel prefixes presented in the resultative I of the transitive verbs have been marked as version markers in origin. A user of the database will not miss them when searching for version

markers. At the same time, it will be highlighted that in these forms they do not express the function of version nowadays.

- Some forms that do not change the form, but can express various functions in various positions have also been distinguished. For example, first and second person personal pronouns can express the function of nominative, ergative or dative without having changed the form. They have also been marked as functional nominative, ergative, and dative.
- Verb classes and declension types are not marked, although inversive verbs are marked. m-series markers are marked as object markers and v-series markers as subject markers, but, at the same time, it is indicated when the verb is inversive.

It was also suggested that each category should be tagged as a verbal or nominal category, and at the same time, as a derivation or inflexional category.

Thus, the database can be used as a facilities by students and linguists who search for linguistic material without knowing the language.

**ქართული ენის კომპიუტერული დამუშავების
ზოგი პრობლემური საკითხისათვის**
(შესიტყვებათა სინტაქსური მოდელები)

თამარ ვაშაკიძე, ნინო ჯორბენაძე, თეა ტეტელოშვილი
არნ. ჩიქობავას ენათმეცნიერების ინსტიტუტი
ninoaka55@yahoo.com

ქართული ენის სრულყოფილი სინტაქსური მოდელირება უპირველეს ყოვლისა საჭიროებს შესიტყვებათა ტიპების ანალიზს, რაც პირველ ეტაპზე გულისხმობს სინტაქსური მექანიზმის დადგენას ზმნა-შემასმენელსა (ძირითად კოორდინატსა) და მასთან დაკავშირებულ სახელებს (მორფოლოგიურ სუბიექტსა და მორფოლოგიურ ობიექტებს) შორის. თანამედროვე ქართულში შესიტყვების ზმნური კოორდინატი ძირითადად მართულია სახელური კომპონენტების მიერ, რადგანაც ამ უკანასკნელთა (სახელურ კოორდინატთა) სინტაქსური ძალა თითქმის ყოველთვის აისახება მასში (ზმნაში) რეალურად ან პოტენციურად. სხვა საკითხია, თუ როგორ წარმოგვიდგება შესიტყვების ძირითადი კოორდინატი მმართველის როლში; რა სახის სინტაქსური ურთიერთობა მყარდება ზმნურ კოორდინატსა და მასთან დაკავშირებულ სახელებს – მორფოლოგიურ სუბიექტსა და მორფოლოგიურ ობიექტებს – შორის ძირითადი კოორდინატის მიერ მათი (სახელების) მართვის თვალსაზრისით. ამ შემთხვევაში გამოიყოფა მართვის ორი ძირითადი ტიპი: 1. **ძლიერი მართვა**, როცა სახელი თუ სახელები (მორფოლოგიური სუბიექტი და მორფოლოგიური ობიექტები) ზმნისაგან მართულნი არიან განსხვავებული ბრუნვის ფორმით დრო-კილოთა ჯგუფების მიხედვით და 2. **სუსტი მართვა**, როცა ზმნა-შემასმენელი გარკვეული ბრუნვის ფორმით მართავს მასთან დაკავშირებულ სახელსა თუ სახელებს ერთი სერიის ფარგლებში, მაგრამ მორფოლოგიური სუბიექტი და მორფოლოგიური ობიექტები ბრუნვას არ იცვლიან დრო-კილოთა ჯგუფების მიხედვით.

პირველ შემთხვევას (ძლიერ მართვას) განეკუთვნება შემდეგი ტიპები:

- ა) უშენებს მამა შვილს სახლს
(ა)უშენა მამამ შვილს სახლი
(ა)უშენებია მამას შვილისთვის სახლი
- ბ) აშენებს კაცი სახლს
(ა)აშენა კაცმა სახლი
(ა)უშენებია კაცს სახლი
- გ) ჰპატრონობს მამა შვილს
უპატრონა მამამ შვილს
უპატრონებია მამას შვილისთვის
- დ) მოაქვს ბავშვს წიგნი

- მოიტანს ბავშვი წიგნს
- მოიტანა ბავშვმა წიგნი
- მოუტანია ბავშვს წიგნი
- ე) ეთამაშება ბავშვი ძაღლს
- ეთამაშა ბავშვი ძაღლს
- უთამაშია ბავშვს ძაღლთან
- ვ) ცხოვრობს კაცი
- იცხოვრა კაცმა
- უცხოვრია კაცს

მეორე შემთხვევას (სუსტ მართვას) განეკუთვნება შემდეგი ტიპები:

- ა) ჰპირდება მამა შვილს გასეირნებას
- დაჰპირდა მამა შვილს გასეირნებას
- დაჰპირებია მამა შვილს გასეირნებას
- ბ) უწითლდება ლოყა ბავშვს
- გაუწითლდა ლოყა ბავშვს
- გასწითლებია ლოყა ბავშვს
- გ) სძულს ბავშვს ფაფა
- სძულდა ბავშვს ფაფა
- სძულებია ბავშვს ფაფა
- დ) სძინავს ბავშვს
- ეძინა ბავშვს
- სძინებია ბავშვს
- ე) მოდის ბავშვი
- მოვიდა ბავშვი
- მოსულა ბავშვი

ცალკე გამოსაყოფია ისეთი შემთხვევები, როცა ერთსა და იმავე ზმნასთან შეწყობილი მორფოლოგიური სუბიექტი წარმოგვიდგება ხან სახელობითი, ხან მოთხრობითი ბრუნვის ფორმით:

- | | | |
|------------------|-------|------------------|
| ა) იტანჯება კაცი | შდრ.: | ბ) იტანჯება კაცი |
| იტანჯა კაცი | | იტანჯა კაცმა |
| დატანჯულა კაცი | | უტანჯია კაცს |

ან კიდევ:

- | | | |
|------------------------|-------|------------------------|
| ა) იწერება კაცი წერილს | შდრ.: | ბ) იწერება კაცი წერილს |
| მოიწერა კაცი წერილს | | მოიწერა კაცმა წერილი |
| მოუწერია კაცს წერილი | | მოუწერია კაცს წერილი |

ზმნა-შემასმენელთა ჯგუფებში გათვალისწინებული უნდა იყოს ყველა ზემოთ წარმოდგენილი ტიპი (შესაძლებელია ამ ტიპების რაოდენობა კიდევ გაიზარდოს). სახელური და ზმნური კოორდინატების კომპიუტერული სისტემატიზაციის შედეგად მომხმარებელი მნიშვნელოვან ინფორმაციას მიიღებს ამა თუ იმ შესიტყვების სინტაქსური მექანიზმის შესახებ.

On Some Problems of Computer Processing of the Georgian Language (Syntactic Models of Word-Correspondences)

Tamar Vashakidze, Nino Jorbenadze, Tea Teteloshvili

Arn. Chikobava Institute of Linguistics

ninoaka55@yahoo.com

The analysis of the types of word-correspondences can be considered to be essential for a complete syntactic modeling of the Georgian language which, at the first stage of analysis, implies establishment of a syntactical mechanism between a verb-predicate (a basic co-ordinate) and the nouns (a morphological subject and a morphological object) connected with it. In modern Georgian a verbal co-ordinate of a word-correspondence is mostly governed by nominal components, as the power of the latter (of nominal co-ordinates) is, in most cases, reflected in the verb either explicitly or potentially. Another issue is the method of representation of a basic co-ordinate of a word-correspondence as governor in respect of the type of a syntactic relation established between a verbal co-ordinate and the nouns connected with it on the one hand (a morphological subject and a morphological object) and a basic co-ordinate from the viewpoint of their (nouns) governing on the other. In this case, two main types of governing are distinguished: **1 a strong governing**, when a noun or nouns (a morphological subject and a morphological object) are governed by a verb through different case forms according to the groups of tenses and moods and **2 a weak governing**, when a verb-predicate governs a noun or the nouns connected with it within the limits of one series of verbs through a certain case form, but the morphological subject and the morphological object do not change their case according to the groups of tenses and mood of the verb.

The following types belong to the first type (strong governing):

-
- a) ušenebs **mama švils saxls** “A father is building a house for a child”
(a)ušena **mamam švils saxli** “A father built a house to a child”
(a)ušenebia **mamas švilistvis saxli** “A father has built a house for a child”
- b) ašenebs **kaci saxls** “A man is building a house”
(a)ašena **kacma saxli** “A man built a house”
(a)ušenebia **kacs saxli** “A man has built a house”
- c) hpatronobs **mama švils** “A father patronages a child”
upatrna **mamam švils** “A father patronaged a child”
upatronebia **mamas švilistvis** “A father has patronaged a child”
- d) moakvs **bavšvs čigni** “A child is bringing a book”
moičans **bavšvi čigns** “A child will bring a book”
moičana **bavšvma čigni** “A child brought a book”
moučania **bavšvs čigni** “A child has brought a book”
- e) etamašeba **bavšvi zayls** “A child is playing with a dog”
etamaša **bavšvi zayls** “A child played with a dog”
utamašia **bavšvs zayltan** “A child has played with a dog”
- f) cxovrobs **ḳaci** “A man is living”
icxovra **ḳacma** “A man lived”
ucxovria **ḳacs** “A man has lived”

The following types belong to the second case (a weak governing):

- a) hpirdeba **mama švils gaseirnebas** “A father is promising a child to go for a walk”
dahpirda **mama švils gaseirnebas** “A father promised a child to go for a walk”
dahpirebia **mama švils gaseirnebas** “A father has promised a child to go for a walk”
- a) učiṭldeba **bavšvs loqa** “A child’s cheek is getting red”
aučiṭlda **bavšvs loqa** “A child’s cheek got red”
aščiṭlebia **bavšvs loqa** “A child’s cheek has got red”
- c) sžuls **bavšvs papa** “A child hates porridge”
sžulda **bavšvs papa** “A child hated porridge”
sžulebia **bavšvs papa** “A child has hated porridge”
- d) sžinavs **bavšvs** “A child is sleeping”
ežina **bavšvs** “A child slept”
sžinebia **bavšvs** “A child has slept”
- e) modis **bavšvi** “A child is coming”
movida **bavšvi** “A child came”
mosula **bavšvi** “A child has come”

The cases must be separated when a morphological subject combined with a same verb is represented sometime in a nominative and sometimes in an ergative case form:

- a) itanžeba **kaci** "A man is tormenting"
daitanža **kaci** "A man tormented"
datanžula **kaci** "A man has tormented"
- cf.: b) itanžeba **kaci** "A man is tormenting"
itanža **kacma** "A man tormented"
utanžia **kacs** "A man has tormented"

or

- a) içereba **ḳaci çerils** "A man is writing a letter" cf.: icereba **ḳaci çerils**
"A man is writing a letter"
moiçera **ḳaci çerils** "A man wrote a letter" moiçera **ḳacma çerili**
"A man wrote a letter"
mouçeria **ḳacs çerili** "A man has written a letter" mouçeri **ḳacs çerili**
"A man has written a letter"

All the above mentioned types of verb-predicates (an increase in number of these types is also possible) must be taken into account in groups of verb-predicates. As a result of a computer systemisation of nominal and verbal co-ordinates the user will obtain important information about a syntactic mechanism of word-correspondences.

ქართული ენის რიცხვითი სახელების კომპიუტერული რეალიზაცია¹

ლიანა ლორთქიფანიძე

სსიპ არჩილ ელიაშვილის მართვის სისტემების ინსტიტუტი
L_Lordkipanidze@yahoo.com

წარმოგიდგენთ ქართული რიცხვითი სახელების კომპიუტერულ რეალიზაციას ენის მოდელირების ექსპერტულ სისტემაში – MESLM. სისტემა რიცხვითი სახელების ფლექსიას ასახავს მარკერებით, რომლებიც აღწერენ რიცხვითი სახელის სახეს და მორფოლოგიურ მახასიათებლებს. იგი ორმიმართულებიანია და შეიძლება გამოყენებული იყოს რომელიმე ბრუნვაში მდგარი რთული რიცხვითი სახელის ანალიზისათვის (მაგ., *ას ოცდათხუთმეტმა*), ან მარკერებით და ციფრებით ასახული შესაბამისი გამოსახულების (მაგალითად, 135+Sg+Gen) გენერირებისათვის.

ქართულ ენაში ციფრების ასახვა რიცხვით სახელებში გაცილებით უფრო რთულია, ვიდრე შესაბამისი გარდაქმნები სხვა ენებში, მაგ., ინგლისურში, ვინაიდან ქართულში რთული რიცხვითი სახელები იბრუნვის, რიცხვების გამოხატვა კი ხდება ოცობით-ათობითი შერეული სისტემით.

ქართულში, ისევე როგორც ბევრ ენაში, რიცხვითი სახელები იყოფა რაოდენობით, რიგობით და წილობით ჯგუფებად. ყველა რაოდენობითი რიცხვითი სახელი ერთიდან ათის ჩათვლით მარტივი რიცხვითი სახელია: 0-*ნული*, 1-*ერთი*, 2-*ორი*, 3-*სამი*, 4-*ოთხი*, 5-*ხუთი*, 6-*ექსი*, 7-*შვიდი*, 8-*რვა*, 9-*ცხრა*, 10-*ათი*. აგრეთვე მარტივია რიცხვითი სახელები: 20-*ოცი*, 100-*ასი*, 1,000,000-*მილიონი*, 1,000,000,000-*მილიარდი*, 1,000,000,000,000-*ტრილიონი*.

თერთმეტიდან ცხრამეტის ჩათვლით ყველა რიცხვითი სახელი არის რთული. ჩვენ შეგვიძლია გამოვიყენოთ არითმეტიკული “ბაზისი” რიცხვითი სახელის გამოსახულების კონსტრუქციისათვის. თერთმეტიდან ცხრამეტის ჩათვლით ქართული რიცხვითი სახელებისათვის იგულისხმება სიდიდე $B=10$ ისე, რომ რიცხვითი სახელების კონსტრუქცია ხდება $[B \cdot n_1 + n_2 + n_1 \cdot \text{მეტი}]$ ფორმულის მიხედვით, ანუ თუ ჩვენ გვაქვს რიცხვი n , მაშინ “ბაზისი” – $B=ათი$ (10), არითმეტიკული მნიშვნელობა $n_1 = n/10$ (აქ და შემდეგში ჩვენ ვიყენებთ მთელი სახის გაყოფას); $n_2 = n \% 10^2$ (1 -დან

¹ აღნიშნული პროექტი განხორციელდა სსიპ – ქართველოლოგიის, ჰუმანიტარული და სოციალური მეცნიერებების ფონდში (რუსთაველის ფონდი) მოპოვებული გრანტის მეშვეობით (გრანტი № A-36-09). წინამდებარე პუბლიკაციაში გამოთქმული ნებისმიერი აზრი ეკუთვნის ავტორს და შესაძლოა არ ასახავდეს სსიპ - ქართველოლოგიის, ჰუმანიტარული და სოციალური მეცნიერებების ფონდის (რუსთაველის ფონდი) შეხედულებებს.

² ჩვენ გამოვიყენებთ % ნიშანს (x/y) განაყოფის ნაშთის (r) აღსანიშნავად. ასე რომ $r = x \% y$ და მაგალითად, როდესაც $x=18$ და $y=10$, მაშინ $r=18 \% 10=8$.

9'-მდე არითმეტიკული მნიშვნელობებისათვის), პლუს მორფემა *მეტი*. მაგალითად: რიცხვ '18'-ს აქვს შემდეგი სტრუქტურა – [ათი+რვა+მეტი] და გარკვეული მორფოლოგიური ალტერნაციების შემდეგ ის აისახება გამოსახულებაში *თვრამეტი*.

ქართულში რიცხვები 20-ის შემდეგ 99-ის ჩათვლით გამოიხატება ოცობით სისტემაში. მომდევნო "ბაზისად" ჩვენ ვიყენებთ სიდიდეს $T=20$ და ვგულისხმობთ T -ს ისე, რომ რიცხვითი სახელების გამოსატვა აიკება შემდეგი გამოსახულების მიხედვით:

$$[[n_1 \cdot T] + n_3 \cdot [\text{და} + [n_2 \cdot B] + n_4 + [n_2 \cdot \text{მეტი}]] \tag{1}$$

ამ სქემის მიხედვით n რიცხვისთვის 1 არის არასავალდებულო მამრავლი და $T=20 = \text{ოცა}$, $B=10 = \text{ათა}$, $n_1 = n/20$; $n_2 = (n \% T) / B$; $n_3 = ((40 - n \% T) / 20) \% 2$; $n_4 = n \% B$.

ასე რომ, რიცხვისთვის $n=57$ არითმეტიკულად ჩვენ გვაქვს $n_1 = 57/20 = 2$; $n_2 = (57 \% 20) / 10 = 1$; $n_3 = ((40 - 17) / 20) \% 2 = 1$; $n_4 = 57 \% 10 = 7$ და რიცხვი 57 აისახება რიცხვით სახელში '57' => $[[2 \cdot T] + 1 \cdot [\text{და} + [1 \cdot B] + 7 + [1 \cdot \text{მეტი}]] = > [[\text{ორი} \cdot \text{ოცი}] + 1 \cdot [\text{და} + \text{ათი}] + \text{შვიდი} + \text{მეტი}]$

=> *ორმოცდაჩვიდმეტი*.

რიცხვისთვის $n=87$ არითმეტიკულად ჩვენ გვაქვს $n_1 = 87/20 = 4$; $n_2 = (87 \% 20) / 10 = 0$; $n_3 = ((40 - 7) / 20) \% 2 = 1$; $n_4 = 87 \% 10 = 7$ და რიცხვი 87 აისახება რიცხვით სახელში

'87' => $[[4 \cdot T] + 1 \cdot [\text{და} + [0 \cdot B] + 7 + [0 \cdot \text{მეტი}]] = > [[\text{ოთხი} \cdot \text{ოცი}] + 1 \cdot [\text{და} + \text{შვიდი}] = > \text{ოთხმოცდაშვიდი}$

რიცხვისთვის $n=80$ არითმეტიკულად ჩვენ გვაქვს $n_1 = 80/20 = 4$; $n_2 = (80 \% 20) / 10 = 0$; $n_3 = ((40 - 0) / 20) \% 2 = 0$; $n_4 = 80 \% 10 = 0$ და რიცხვი 80 აისახება რიცხვით სახელში

'80' => $[[4 \cdot T] + 0 \cdot [\text{და} + [0 \cdot B] + 0 + [0 \cdot \text{მეტი}]] = > [[\text{ოთხი} \cdot \text{ოცი}] + 0 \cdot [\text{და} + 0 + 0]] = > \text{ოთხმოცი}$

ქართულში რიცხვები 100-ის შემდეგ გამოიხატება ოცობით-ათობითში, მაგრამ უმრავლეს ენებში რიცხვითი სახელების გამოსატვის ინტერპრეტაცია სქემატურად მსგავსია და აისახება შემდეგი ფორმულით: $V = [M \cdot F + R]$. რიცხვითი სახელის რიცხვითი სიდიდე V (Value) მიიღება სამრავლის M (Multiplicand) მამრავლზე F (Factor) გადამრავლებით და R (Remainder) ნაშთის მიმატებით. ამ რეკურსიულ სტრუქტურაში M და R კომპონენტები თვითონვე შეიძლება იყვნენ რთული რიცხვითი სახელები, რომლებიც ექვემდებარებიან შემდეგ შეზღუდვებს:

1. $R = V \% F$;
2. $M = V / F$;
3. რიცხვ ასამდე სამრავლი მრავლდება მამრავლზე '20'; ასის შემდეგ – მამრავლზე '10' და ათასის შემდეგ – მამრავლზე '1000';
4. M და R არასავალდებულო კომპონენტებია. თუ გამოტოვებულია M , იგულისხმება რომ ის უდრის 1-ს, თუ გამოტოვებულია R , იგულისხმება რომ ის უდრის 0-ს;
5. თუ R -ის (ან M -ის) რიცხვითი სიდიდე არის 21-დან 99-ის ჩათვლით, მაშინ რიცხვითი სახელები (1) გამოსახულებაზე დაყრდნობით იქნება შედგენილი.

მაგალითად, *ოცდათხუთმეტი ათას შვიდას ოთხმოცდასამისთვის* რთულ რიცხვით სახელს აქვს შემდეგი სტრუქტურა:

[[*ოცი*]+*და*+[*ათი*+*ხუთი*+*მეტი*]]·*ათი*·*ასი*+[*შვიდი*·*ასი*]+[*ოთხი*·*ოცი*]+*და*+*სამი*]]
და რიცხვებში, 35,783=35·1000+7·100+83= [[1·20+[10+5]]·1,000+[7·100]+[4·20]+3]]].

რთული რიცხვითი სახელის მხოლოდ ბოლო ნაწილი ექვემდებარება ბრუნებას. ასევე ხდება მრავლობითის წარმოებისას. მაგალითად: '139+Pl+Gen'=ას ოცდაცხრამეტების ('Pl'=ებ და 'Gen'=ის). რიცხვი 1-ის გარდა ('1st+Pl+Dat'=პირველებს), რიგობითობა მარკირებულია ცირკუმფიქსით - *მე* -*გ*.

მაგალითად: '4000th+Sg+Nom'=მეოთხიათასე, '239th+Sg+Nom'=ორასოცდამეცხრამეტე, '90,000th+Pl+Gen'=ოთხმოცდამეათასეების.

წილობითობა მარკირებულია სუფიქსით, რომელიც წინ უსწრებს რიცხვისა და ბრუნვის მარკერებს. მაგალითად: '/1,0004000th + Sg + Nom' = მილიონმეოთხიათასედი, '/139th+Sg+Na'=ას ოცდამეცხრამეტედი. რთული რიცხვითი სახელები იწერება დანაწევრებით - *სამოცდაცამეტი ათას ხუთას ერთი*='73501+Sg+Nom'. რიგობითი და წილობითი რიცხვითი სახელები კი ყოველთვის იწერება როგორც ერთი სიტყვა.

MESLM სისტემაში მორფოლოგიური ლექსიკონი დაიყო კანონიკურ და წესების ლექსიკონად. კანონიკური ლექსიკონი (CaDic) არის მონაცემთა ბაზის ფაილი, რომელშიც შედის სიტყვის კანონიკურ ფორმასა და მისი მორფოტაქტიკური წესების მარკერებს შორის შესატყვისობები. მორფოტაქტიკური წესები აღწერილია ორ სალექსიკონო მონაცემთა ბაზაში: ფორმაწარმოების (InRuLex) და სიტყვაწარმოების (DeRuLex) წესების ლექსიკონში. სახელების ბრუნება, ზმნების უღლება, პროდუქტიული სიტყვაწარმოება და შერწყმა (სახელების გაერთიანება) უზრუნველყოფილია დამატებითი ლექსიკონებით. ესენია: ფორმაწარმოების (InFeLex) და სიტყვაწარმოების (DeFeLex) გრამატიკული მახასიათებლების ლექსიკონები. მაგალითად, ქართულ სიტყვაფორმას *მერვესი* CaDic-ში შეესაბამება ლემა *რვა*, მისი ფორმაწარმოების წესი (+*lsm*) და სიტყვაწარმოების წესი (*მე^{pm}*+, +*გsm*), ხოლო მახასიათებლების ლექსიკონებში, შესაბამისად, სიტყვაწარმოების 8+Ord და ფორმაწარმოების მახასიათებლები - Sg+Gen. ასე რომ: *მერვესი*=*მე*+*რვა*+*ესი*= '8'+Ord+Sg+Gen. რიცხვითი სახელის ფორმირების ინფორმაციის (1) ფორმულაში კოდირებით და ზემოთ ჩამოთვლილი შეზღუდვებით *ათასხუთასოთხმოცდამეთვრამეტეებმა*-ს ასახვა გამოსახულებაში - '1598'+Ord+Pl+Erg და მდგომარეობათა სასრული რაოდენობის მარტივ ოპერატორებში მისი რეალიზება შემდეგნაირადაა შესაძლებელი:

ათ=>10; *ას*=>100; *ხუთ*=>5; *ოთხ*=>4; *ოცი*=>20; *რვა*=>8; *მე^{pm}*+*გsm*=>+Ord; +*ებsm*=>+Pl; +*მაsm*=>+Erg, ანუ *ათასხუთასოთხმოცდამეთვრამეტეებმა* =>

[[*ათი*·*ასი*]+ [*ხუთი*·*ასი*]+[*ოთხი*·*ოცი*]+*და*+*მე^{pm}*]+*ათი*+*რვა*+*მეტე*]+*ებsm*+*ებsm*+*მაsm*]=>

[[10·100]+[5·100]+[4·20]+[10+8]+Ord+Pl+Erg]=>[[1000]+[500]+[80]+[18]+Ord+Pl+Erg]=> '1598'+Ord+Pl+Erg

Computational Implementation of the Georgian Numerals¹

Liana Lortkipanidze

LEPL Archil Eliashvili Institute of Control Systems

L_Lortkipanidze@yahoo.com

This paper presents a method of computational implementation of the Georgian numeral system in the Multilingual Expert System of Language Modeling (MESLM) [1]. The system is bidirectional and represents inflection of numerals with markers that indicate their ordinality and morphological features. It can be used to analyse complex numerals in any of the cases of the system (for example, *ახ ოცდათხუთმეტი*) and to generate their corresponding expression represented with markers and numbers (such as 135+Sg+Gen).

In the Georgian language the representation of numbers into numerals is much more complicated than similar transformations in languages such as English. This is because Georgian complex numerals are inflected and numbers are expressed by a vigesimal–decimal mixed system.

Like most languages, numerals in Georgian are classified as Cardinals, Ordinals, and Fractions groups. All cardinal numerals from one to ten are simple numerals: 0 – *ნული*, 1 – *ერთი*, 2 – *ორი*, 3 – *სამი*, 4 – *ოთხი*, 5 – *ხუთი*, 6 – *ექვსი*, 7 – *სვიდი*, 8 – *რვა*, 9 – *ცხრა*, 10 – *ათი*. 20 – *ოცი*, 100 – *ასი*, 1,000,000 – *მილიონი*, 1,000,000,000 – *მილიარდი*, 1,000,000,000,000 – *ტრილიონი* are simple numerals too.

All numerals from *eleven* to *nineteen* are complex numerals. Arguably, we can use the arithmetical “base” to construct numeral expressions. In the Georgian numeral system for numbers from eleven including nineteen we can use the value $B = 10$, so that numeral expressions are constructed according to the formula $[B \cdot n_1 + n_2 + n_1 \cdot \text{მეტი}]$, i.e. if we have a number n , then the “Base” – $B = \text{ათი}$ (‘10’) arithmetical value $n_1 = n/10$ (here and further, we use a whole division); $n_2 = n \% 10^2$ (for arithmetical values from ‘1’ to ‘9’), plus the morpheme *მეტი* (meaning ‘more’). For example: number ‘18’ has the structure – $[\text{ათი} + \text{რვა} + \text{მეტი}]$ and after some morphophonological alternations is mapped to *თვრამეტი*.

¹ The designated project has been fulfilled by financial support from the Foundation For Georgian Studies, Humanities and Social Sciences (Rustaveli Foundation) (Grant № A-36-09). Any idea in this publication is that of the author and may not represent the opinion of Foundation For Georgian Studies, Humanities and Social Sciences (Rustaveli Foundation) itself.

² We use the division sign % for remainders (r) of divisions (x/y). So $r = x \% y$. For example when $x = 57$ and $y = 20$, then $r = 57 \% 20 = 17$.

In Georgian the numbers from 20 up to 99 are expressed by a vigesimal system. We use the value T = 20 as the next “Base” so that numeral expressions are constructed according to the formula:

$$[[n_1 \cdot T] + n_3 \cdot [\text{ღა} + [n_2 \cdot B] + n_4 + [n_2 \cdot \text{მეტი}]]] \tag{1}$$

According to this scheme, 1 is a needless multiplier for number n and T = ‘20’ = ოცო; B = ‘10’ = ათო; $n_1 = n/20$; $n_2 = (n \% T)/B$; $n_3 = ((40 - n \% T)/20) \% 2$; $n_4 = n \% B$.

Thus for the number n = 57 arithmetically we have $n_1 = 57/20 = 2$; $n_2 = (57 \% 20)/10 = 1$; $n_3 = ((40 - 17)/20) \% 2 = 1$; $n_4 = 57 \% 10 = 7$ and the number 57 is represented into the numeral ‘57’ = $\> [[2 \cdot T] + 1 \cdot [\text{ღა} + [1 \cdot B] + 7 + [1 \cdot \text{მეტი}]] = \> [[\text{ორი} \cdot \text{მოცი}] + 1 \cdot [\text{ღა} + [\text{ათი}] + \text{შვიდი} + [\text{მეტი}]]]$

=> ორმოცდახვიდმეტი.

For the number n = 87 arithmetically we have $n_1 = 87/20 = 4$; $n_2 = (87 \% 20)/10 = 0$; $n_3 = ((40 - 7)/20) \% 2 = 1$; $n_4 = 87 \% 10 = 7$ and the number 87 is represented into the numeral ‘87’ = $\> [[4 \cdot T] + 1 \cdot [\text{ღა} + [0 \cdot B] + 7 + [0 \cdot \text{მეტი}]] = \> [[\text{ორი} \cdot \text{მოცი}] + 1 \cdot [\text{ღა} + \text{შვიდი}]] = \> \text{ოთხმოცდაშვიდი}$.

For the number n = 80 arithmetically we have $n_1 = 80/20 = 4$; $n_2 = (80 \% 20)/10 = 0$; $n_3 = ((40 - 0)/20) \% 2 = 0$; $n_4 = 80 \% 10 = 0$ and the number 80 is represented into the numeral ‘80’ = $\> [[4 \cdot T] + 0 \cdot [\text{ღა} + [0 \cdot B] + 0 + [0 \cdot \text{მეტი}]] = \> [[\text{ოთხი} \cdot \text{მოცი}] + 0 \cdot [\text{ღა} + 0 + 0 + 0]] = \> \text{ოთხმოცი}$.

In Georgian Numbers above 100 are expressed vigesimal–decimally but schematically the interpretation of numeral expressions is very similar in most languages – [M • F + R]. The Value (V) of the numeral is obtained by multiplying M (Multiplicand) by F (Factor) and adding R (Remainder) to the result. In this recursive structure the M and R components may themselves be complex numerals. But they obey the following restrictions:

1. $R = V \% F$;
2. $M = V/F$;
3. Up to 100 the Multiplicands multiply by the factor 20; after 100 by a factor of 10 and after 1000 - by a factor of a thousand;
4. M and R are optional components. A missing M is equivalent to 1, a missing R is interpreted as 0;
5. if the value of R (or M) is from 2 to 99, the numerals are constructed according to the pattern.

For example, the complex numeral ოცდათხუთმეტი ათას შვიდას ოთხმოცდასამი ‘35,783’ has the structure:

[[[*ოცი*]+*და*+[*ათი*+*ხუთი*+*მეტი*]] ·*ათი ასი*]+[[*შვიდი ასი*]+[*ოთხი ოცი*]+*და*+*სამი*];

in numbers, [[1·20+[10+5]] ·1,000+[[7·100]+[4·20]+3]]= >35,783.

Only the last part of a complex numeral gets marks of number and case. For example: ‘139 + Pl + Gen’ = *ასოცდაცხრამეტების*, where ‘Pl’ = *ებ* and ‘Gen’ = *ის*. Ordinality is marked by the circumfix *მე-გ*, except of number 1 (‘1st+Pl+Dat’=*პირველებს*).

For example: ‘4000th+Sg+Nom’ = *მეოთხიათასე*, ‘239th+Sg+Nom’ = *ორასოცდამეცხრამეტე*, ‘90,000th+Pl+Gen’ = *ოთხმოცდამეათასების*. Fractionality is marked by adding the suffix *გდო* after the ordinal numeral and goes before number and case markers. For example: ‘/1,0004000th+Sg+Nom’=*მილიონმეოთხიათასედი*, ‘/139th+Sg+Na’=*ასოცდამეცხრამეტეგდო*. Complex cardinal numerals are written separately *სამოცდაცამეტი ათას ხუთას ერთი* = ‘73501+Sg+Nom’. But ordinal and fractional numerals always are written as one word.

In the MESLM system the morphological dictionary has been divided into a canonical dictionary and two lexicons of rules. The canonical dictionary (CaDic) is a database file containing correspondences between canonical forms of the word and markers of morphotactic rules. The rules of morphotactics have been described in the databases of two lexicons: inflection (InRuLex) and derivation (DeRuLex) rule lexicons. Noun declination, verb conjugation, comparison of adjectives, productive derivation and compounding are provided with support of the continuation lexicons: inflection (InFeLex) and derivation (DeFeLex) grammatical features lexicons. For example, the Georgian word form *მერვეს* corresponds to the CaDic with lemma *რვა*, its inflexion rule +*ს^S*, derivation rule *მე^P*+, +*ე^S* and correspondingly inflection Sg+Gen and 8+Ord – derivation features. So *მერვეს*=*მე+რე+ეს*=‘8’+Ord+Sg+Gen. The complex numeral formations are built upon other lexemes. Coding of the numeral formation information by formula (1) and by considered limitations allows us to construct the form *ათასხუთასოთხმოცდამეთვრამეტეებმა* to ‘1598’+Ord+Pl+Erg and to realise into simple finite-state operations:

ათ=>10; *ას*=>100; *ხუთ*=>5; *ოთხ*=>4; *ოც*=>20; *რვა*=>8; *მე^P*+*ე^S*=>+Ord;; +*ებ^S*=>+Pl; +*მა^S*=>+Erg, so

ათასხუთასოთხმოცდამეთვრამეტეებმა=>

[[*ათი* ·*ასი*]+ [*ხუთი* ·*ასი*]+[*ოთხი* ·*ოცი*]+*და*+*მე^P*+[*ათი*+*რვა*+*მეტი*]+*ებ^S*+*მა^S*]=>

[[10·100]+[5·100]+[4·20]+[10+8]+Ord+Pl+Erg]=>[[1000]+[500]+[80]+[18]+Ord+Pl+Erg]

=>‘1598’+Ord+Pl+Erg

ქართული ზმნის პარადიგმათა ელექტრონული ვერსია და მისი მნიშვნელობა უცხოელთათვის ქართული ენის სწავლების პროცესში

ნანა შავთვალაძე, ნინო შარაშენიძე
თბილისის სახელმწიფო უნივერსიტეტი

ზმნა ქართული ენის ურთულესი და უმთავრესი ნაწილია. ზმნურ ფორმათა გაანალიზებისა და სწორად გამოყენების გარეშე მარტივი კომუნიკაციის აგებაც კი შეუძლებელია ენის შემსწავლელთათვის. დღეს აქტუალურია ისეთი ელექტრონული პროგრამების შექმნა, რომლებიც მათ პრობლემების დაძლევაში დაეხმარება. ერთ-ერთ ასეთ მნიშვნელოვან სისტემად გვევლინება ზმნურ პარადიგმათა ელექტრონული ვერსიის შექმნა. მას საფუძვლად დაედება, ერთი მხრივ, კომუნიკაციისას ზმნურ ფორმათა გამოყენების სისწორული ანალიზი, მეორე მხრივ, გათვალისწინებული იქნება ზმნათა ფორმობრივი კონსტრუქცია და სტრუქტურა.

კონსტრუქციისა და სტრუქტურის მიხედვით ზმნები დაყოფილია ოთხ ტიპად და 14 ქვედანაყოფად. პირველ ტიპში ექვს ნაწილადაა წარმოდგენილი ყველა გარდამავალი და საშუალო-მოქმედებითი ზმნა. ზმნური ფორმები დალაგებულია თემის ნიშანთა მიხედვით: -ობ, -ებ, -ამ, -ავ, -ი და უთემისნიშნო.

<i>The conjugation of verbs is divided into four types and 14 parts</i>			
I ტიპი	II ტიპი	III ტიპი	IV ტიპი
№1 --stem-ობ	№7 მ--stem-	№9 --stem-ები	№14 irregulars
№2 --stem-ებ	გ--stem-	№10 -stem-დები	
№3 --stem-ამ	ბ,პ,ს--stem-	№11 ი--stem-ები	
№4 --stem-ავ	№8 მი--stem-	№12 ე--stem-ები	
№5 --stem-ი	ვი--stem-	№13 --stem--ია	
№6 --stem-ო	უ--stem-		

II ტიპში გაერთიანებულია ინვერსიული, მანის რიგის, ზმნები. ამ რიგში ორი, მე-7 და მე-8, ქვეტიპია გაერთიანებული.

III ტიპის ზმნებში თავმოყრილია ვნებითი გვარის ზმნები. მესამე ტიპში ხუთი ნაწილია. დინამიკურ ვნებითს ეთმობა ოთხი (მე-9 – მე-12), ხოლო სტატიკურს ერთი – მე-13 ნაწილი.

IV ტიპში შედის ყველა სახის თავისებური ზმნა. მასში მხოლოდ ერთი, მე-14, ქვეტიპია გამოყოფილი.

პროგრამა მოიცავს ზმნათა **საძიებო სისტემას**, რომელსაც ასევე დაერთვება რამდენიმე ბმული:

ზმნის სწავლებისას და მისი ანალიზისას დიდი მნიშვნელობა ენიჭება ენის შემსწავლელისათვის **პარადიგმის** მიწოდებას. პროგრამა გულისხმობს ამ პარადიგმათა სრული სურათის მიწოდებას, მათი სემანტიკური მნიშვნელობებისა და ფუნქციების გაცნობას. ბაზა იქნება ორენოვანი, თუმცა ენათა რაოდენობა შესაძლებელია გაიზარდოს მოთხოვნის შესაბამისად და პროგრამული სრულყოფის კვალდაკვალ.

ზმნურ ფორმათა ათვისებისათვის დიდი მნიშვნელობა ენიჭება კონსტრუქციის ანალიზს. ქართული ზმნის თავისებურებათა გათვალისწინებით **წინადადება** არის ის საშუალება, რომელიც კარგად წარმოაჩენს ზმნით გამოსატული მოქმედების კონკრეტულ მნიშვნელობას და ამავე დროს ენის შემსწავლელს გააანალიზებინებს წინადადების სხვა წევრების ზმნასთან შეთავსება-შეწყობის წესებს.

მნიშვნელოვანია **რთულ სინტაქსურ კონსტრუქციებში** ზმნათა შეფარდებებისა და დროთა თანმიმდევრობის ამსახველი მასალის მოწოდება. კონსტრუქციების გაანალიზება ხელს შეუწყობს ზმნურ ფორმათა მიმართებების ათვისებას.

ცალკე ბმულში მოთავსებული იქნება გაანალიზებულ ზმნურ ფორმებზე აგებული **ტექსტები**, სპეციალურად შექმნილი თუ ავთენტური.

ქართული ზმნის პარადიგმათა ელექტრონული ვერსია ხელს შეუწყობს ენის ათვისებას და გამოყენებული იქნება როგორც დისტანციური, ისე კონტაქტური სწავლების პროცესში.

The Electronic Version of Georgian Verb Paradigms and its Importance in Teaching the Georgian Language to Foreigners

Nana Shavtvaladze, Nino Sharashenidze

Tbilisi State University

The verb is the major and at the same time the most complicated word form of the Georgian language. If not analysed and used correctly the verbal forms may become a deterrent to the simplest forms of communication. Clearly, there is a great need for the creation of electronic programs which will support the language learners to overcome the difficulties in the process of correct usage of verb-forms. One such system is the electronic version of verbal paradigms. This programme is based, on the one hand, on frequency analysis of the verbal forms in the process of communication and on the other hand, on the specificity of construction of the verb-forms. Their structure will also be considered.

According to the types of verbal constructions and their structure, verbs are divided into four types and fourteen subtypes. The first type of verb includes all transient and medium active verbs which are presented in six subparts. The verbal forms are grouped according to the verbal endings of the stem: **-ob**, **-eb**, **-am**, **-av**, **-i**, and ones without a stem verbal ending sign.

<i>The conjugation of verbs is divided into four types and fourteen parts</i>			
I ტიპი	II ტიპი	III ტიპი	IV ტიპი
№1 --stem- ob	№7 m --stem-	№9 --stem- ebi	№14 irregulars
№2 --stem- eb	g --stem-	№10 -stem- debi	
№3 --stem- am	o,h,s -stem-	№11 i --stem- ebi	
№4 --stem- av	№8 mi --stem-	№12 e --stem- ebi	
№5 --stem- i	gi --stem-	№13 --stem- ia	
№6 --stem- o	u --stem-		

The second type includes inversion verbs and the verbs of “MANI” rank. These forms of the objective range fall into the 7th and 8th subparts.

The third type includes those verbs which possess the passive voice. This type is further subdivided into five parts. Four parts (from 9th to 12th) choose the dynamic passive and only one (13th) uses static passive.

The fourth type which only has one subtype (14th subtype) unites all kinds of irregular verbs.

The program contains several verb search systems, which will be tied to several links.

It is very important to provide learners the verb paradigm while teaching and analysing the verb. The programme plans inclusion of the whole range of these paradigms and exploration of their semantics and functions. Initially, the programme will be bilingual, though the number of languages may be increased depending on the demand and the programmes's development.

კვაზიპარადიგმატული მოდელი უდეტრების მონაცემთა ბაზის წარმოდგენისათვის (საკითხის დასმის სახით)

ლევან ჩხაიძე

საქართველოს ტექნიკური უნივერსიტეტი

levan@dr.com

ქართულ ენაში ყველა შესაძლო სიტყვაფორმის მონაცემთა ბაზის შექმნისათვის უადრესად ნაყოფიერი აღმოჩნდა პარადიგმატული (პროდუქციული) მიდგომა, რომლის მიხედვითაც ნებისმიერი ზმნური ან სახელური მართებული სიტყვაფორმა მიიღება ფუძისა და პრეფიქს-სუფიქსების შესაბამისი კომბინაციის პროდუქციული წესების საშუალებით. პარადიგმის მიხედვით, სიტყვაფორმების ტოტალური დაგროვების მეთოდისგანსხვავებით, ეს გზა იძლევა შეუდარებლად უფრო კომპაქტურ შედეგს, მაგრამ მოითხოვს პარადიგმის ნებისმიერი წევრის აგების ზუსტ და ადეკვატურ ფორმალიზაციას. ეს განსაკუთრებით მნიშვნელოვანია აგლუტინაციური ენებისათვის.

ამ მიდგომის მიღმა რჩება ე.წ. ფორმაუცვლელი სიტყვების (უდეტრების) მონაცემთა ბაზაში წარმოდგენის საკითხი, რადგან, არსებული გრამატიკული კონცეფციის თანახმად, მათ არა აქვთ პარადიგმა – ფორმათა ერთობლიობა, რომელიც რომელიმე პროდუქციული წესის მიხედვით გრამატიკულადაა დასაბუთებული. ამ თვალსაზრისიდან გამომდინარე, უდეტრების მონაცემთა ბაზის შექმნა შესაძლებელია მხოლოდ დაგროვების პრინციპით, რაც ეწინააღმდეგება მონაცემთა ბაზის დანარჩენი ნაწილის წარმოდგენის პრინციპს. ამასთან, დაგროვების მეთოდი, მისთვის არსებული მონაცემთა ბაზის აქტუალიზაციის სქემიდან გამომდინარე, ნაკლებად ეფექტურია ლექსიკური სისრულის (ე.ი. ენაში პოტენციურად მართებული ყველა სიტყვაფორმის ასახვის) მისაღწევად.

როგორც ცხადყო ი. ჯანაშიას მიერ უდეტრების 10000 სიტყვაფორმაზე ჩატარებულმა ექსპერიმენტმა, მათი აბსოლუტური უმრავლესობა (პირველ რიგში, ზმნისართები) იყოფა ერთი და იმავე სტრუქტურის ჯგუფებად, რომლებიც შეიცავენ ერთმანეთისგან მხოლოდ ნაწილაკებით განსხვავებულ სიტყვაფორმებს, მაგალითად:

შორს, შორსა, შორსაა, შორსაც, შორსაცაა, შორსეე, შორსლა.

ნიშანდობლივია, რომ ნ. ჩომსკის თეორიის თვალსაზრისით, ამგვარ ჯგუფში შემავალი სიტყვაფორმები თავისუფლად შეიძლება ჩაითვალოს სულ რამდენიმე პროდუქციული წესის საშუალებით ერთი ფორმიდან მიღებულ ფორმად. აქ ბუნებრივად ჩნდება ფორმალური ანალოგია პარადიგმის ცნებასთან, რომელშიც „კვაზი-ფუძის“ ფორმალურ როლს თამაშობს ნაწილაკმოკლებული სიტყვაფორმა (ამ შემთხვევაში **შორს**), ხოლო „კვაზიდაბოლოების“ ფორმალურ როლს – ესა თუ ის ნაწილაკი, ზოგჯერ – **“არის”** ზმნის შეკვეცილი ფორმა, ან მათი კომბინაცია (შე-

საბამისად, -ა, -აა, -აც, -აცაა, -ვე, -ლა). ეს ანალოგია კიდევ უფრო თვალსაჩინო ხდება ექსპერიმენტში დადასტურებული პროდუქციული წესების შესაძლო კომბინაციების (“კვაზიფლექსური ფორმაწარმოების” ფორმალური ტიპების) შედარებით უმნიშვნელო რაოდენობით (რაც, ფაქტობრივად, შეპირობებულია სწორედ ბრუნვის ფორმათა არარსებობით – უდეტრების გრამატიკული ფორმაუცვლელობით). მაგალითად, ზმნისართისათვის პროდუქციათა წყების ყველაზე გავრცელებული ვარიანტი ასე გამოიყურება:

ადვილად, ადვილად-აა, ადვილად-აც, ადვილად-ვე.

კერძოდ, ხსენებულ მასალაში, თუ არ ჩავთვლით ერთფორმიან ჯგუფებს, აღმოჩენილი იყო სულ 12-მდე განსხვავებული სტრუქტურის ჯგუფი.

ტრადიციული გრამატიკული მიდგომის თვალსაზრისით, ნაწილაკების („კლიტიკების“) დართვა არ წარმოადგენს სინტაქსურ მოვლენას, იმიტომ რომ ამ შემთხვევაში სიტყვაფორმების შეცვლა არ იწვევს წინადადების სხვა წევრებთან მათი სინტაქსური ურთიერთობის ხასიათის შეცვლას და ძირითადად გამოიყენება პროსოდიული მიზნებისათვის. ამდენად, განხილული მოვლენის აღწერისას არ უნდა იყოს გამოყენებული ტერმინი „პარადიგმა“ მისი ტრადიციული მნიშვნელობით.

მეორე მხრივ, მიუხედავად გრამატიკული ინტერპრეტაციის არარსებობისა, ერთ ჯგუფში მოხვედრილი სიტყვაფორმების პროდუქციული წესების საშუალებით ერთმანეთთან ფორმალური დაკავშირება მეტად მნიშვნელოვანია მონაცემთა ბაზის სტრუქტურის უნიფიკაციისათვის და, რაც უფრო ფასეულია, მისი სისრულისათვის. ეს მოსაზრება უდავოდ ამართლებს ტერმინ „კვაზიპარადიგმატული მოდელის“ გამოყენებას.

უნდა აღინიშნოს, რომ „კვაზიპარადიგმატული“ მოდელის ელემენტები შეიმჩნევა მონაცემთა ბაზის იმ ფრაგმენტებშიც, რომლებშიც ეჭვს არ იწვევს პარადიგმის არსებობა, ამ ტერმინის ჩვეულებრივი გაგებით. მაგალითად, ფუძე **კაც-** პარადიგმის შემადგენლობაში შედის არა მარტო სიტყვაფორმები: **კაცი, კცმა, კცისათვის**, არამედ: **კაცივე, კცმაც, კცისათვისლა** და ა.შ.

რა თქმა უნდა, საკამათოა, თუ რამდენად მიზანშეწონილი იქნება ქართული ენის ლინგვისტური მოდელის ისეთი გაფართოება, რომელიც განაზოგადებს სინტაქსური კავშირების ცნებას და უზრუნველყოფს „კვაზიპარადიგმატული“ (ე.ი. ცალკე აღებული თვისებების ამსახველი) მოდელის ჩართვას ტრადიციული პარადიგმის ჩარჩოებში. საუკლისხმოა, რომ თეორიის მსგავსი განვითარება ხშირად ხდება მეცნიერების ისტორიაში: საყოველთაოდ ცნობილია გამოჩენილი მოაზროვნის განცხადება წარმოსახვითი რიცხვების შესახებ: „მე ვხედავ მათ, მაგრამ არ მჯერა მათი არსებობისა“.

მით უმეტეს, ეჭვგარეშეა, რომ „კვაზიფლექსური ფორმაწარმოების“ ტიპების უფრო დეტალური შესწავლა, პირველ რიგში, მათი შესაძლო კავშირი „კვაზიფუძის“ ფონეტიკურ სტრუქტურასა და სემანტიკურ თავისებურებებთან, წარმოადგენს რეალურ პრაქტიკულ ინტერესს ლინგვისტურ მონაცემთა ბაზების სრულყოფისათვის. ეს პრობლემა შეიძლება გახდეს შემდგომი ლინგვისტური კვლევის ობიექტი.

Usage of a Quasiparadigmatic Model for the Presentation of a database of Udetres

(Introducing a topic for discussion)

Levan Chkhaidze

Technical University of Georgia

levan@dr.com

For the purpose of creating a database for all the possible word-forms in Georgian the paradigmatic (productional) approach has proved to be fruitful. According to this approach any correct verbal or nominal word-form is received in compliance to the productional rules based on the type of the paradigm and the combinations of the stem, prefixes and suffixes.

In difference from the theory of Total Collection of Word-forms, this approach presents a much more compact scheme for the flexional and especially agglutination languages. However, it demands correct and adequate formalisation for the construction of each member of the paradigm.

The problem of presentation of the so-called “Udetres” (non-changeable word-forms) in the database remains beyond this approach. According to the present grammatical conception they don’t possess a paradigm that is the grammatically adopted entity received by some productional rule. From this point of view, the creation of the data base of the “Udetres” is possible only if the principle of collection is applied, which opposes the presentation principle for the remaining part of

the database. In addition this method, because of the specificity of the actualisation scheme of its database, proves to be less effective in respect of the presentation of all the potential word-forms of the language.

As revealed by the extensive experimental study of 10,000 non-changeable, word-forms conducted by I.Janashia, their absolute majority (principally, adverbs) is divided into groups of similar structures/word-forms which differ only by the ending particles. For instance:

šors, šorsa, šorsaa, šorsac, šorsacaa, šorsve, šorsya.

It is worth noting that from the standpoint of Chomsky theory the word-forms –parts of such groups can be considered to be a form received from one form merely by applying several productional rules. This naturally brings about the formal analogy with the concept of the paradigm where the formal role of the “quasi-stem” is attributed to the word-form used without the particle (in this case *šors*), and the formal role of the “quasi-ending” is played by a particle or in some cases by the contracted form of the verb “არის”:

(to be) or the combination of both (respectively *-a, -aa, -ac, -acaa, -ve, -ya*). This analogy becomes even more visible by the relatively small number of possible combinations of productional rules which can be classed as “quasi-flexional derivation”. This is conditioned by the lack of cases and general grammatical non-changeability of the form typical of Udetres. For instance, the most widely spread variant of the set of productions for the adverb can be represented as follows:

advilad, advilad-aa, advilad-ac, advilad-ve.

More specifically, if the groups characterised by only one form are not taken into consideration, according to the data, the number of the groups characterised by variety of structures will be approximately 12.

According to the point of view of Traditional Grammar the addition of particles is not a syntactical phenomenon as the change of the word-form does not provoke the change of syntactical connections with other members of the sentence and is primarily used for prosodic means.

Consequently, the term “paradigm” should not be applied in its traditional sense to denote the phenomenon discussed above.

On the other hand, in spite of the lack of grammatical interpretation, achievement of the formal connection between the word-forms by members of the group by means of productional rules is of major importance for the unification of the structure of the database and its completeness. Thus the application of the term “quasi pragmatic model” seems to be more appropriate. It must also be noted that the elements of the “quasi-paradigmatic” model are visible in those fragments of the database which possess a traditional paradigm. For example, the paradigm of the stem *კაც* includes not only the following word-forms *კაცი, კაცმა, კაცისთვის* but, also, the word-forms *კაცვე, კაცმაც, კაცისთვისა*.

It is still a matter of discussion how correct it is to broaden the linguistic models of the Georgian language which will certainly broaden the concept of the syntactic connections and provide the inclusion of the “Quasi-paradigmatic” (i.e the model which depicts individual features) model into the frames of the traditional paradigm. Such developments of a theory are quite well-known in the history of the scientific thought: for instance, the following statement of a well-known thinker about imaginary figures is known by everyone: “I see them but I do not believe in their existence”.

Moreover, the more detailed study of the “quasi-flexional” types of word formation in respect of their possible connection with the phonetic structure of the “quasi stem” and specific semantic features presents considerable practical interest for the perfection of linguistic databases. This issue can become the object of further linguistic studies.

სავარაუდო ფორმების ეფექტური ძიების საშუალებები ქართული მართლმწერის ინტერნეტვერსიაში

ლევან ჩხაიძე, დავით ჯაში
Technical University of Georgia

levan@dr.com

იმ ამოცანათა შორის, რომელთა გადაწყვეტაც მართლმწერის შემოწმების თანამედროვე ავტომატიზებული სისტემების (ე.წ. სპელჩეკერების) შესაძლებლობების ნუსხაში შედის, ერთ-ერთი ყველაზე მნიშვნელოვანია სუგესტიური (სავარაუდოდ შეფასებული, სისტემის მიერ ნაგულისხმევი, მოტივირებულად ნაკარნახევი, ოპერატორისთვის ჩაგონებული) მართებული სიტყვაფორმების მოძიება, რომლებიც გარკვეულწილად ახლოსაა გამოსაცნობ სიტყვასთან (უმრავლეს შემთხვევაში – შეცდომის გამო). მართლმწერის ავტომატიზებული შემოწმების გარდა ეს ამოცანა მეტად აქტუალურია სხვა პროგრამული პროდუქტების შესაქმნელადაც, მათ შორის ისეთი რთული სისტემებისათვის, როგორცაა ტექსტურ მონაცემთა ბაზაში ძიების სისტემები, რომლებიც იყენებენ ბუნებრივ ენას.

სუგესტიურ სიტყვაფორმათა ძიების ალგორითმებს აუცილებლად უნდა ჰქონდეს ევრისტიკული ხასიათი, რადგან ისინი ეფუძნება ტექსტში შეცდომის გაჩენის სხვადასხვა ჰიპოთეზას, რომელიც შემუშავებულია შეცდომათა ადამიანური ბუნების გათვალისწინებით. მათ შორის საყოველთაო აღიარება პოვა „მჭიდრო ტექსტუალური სიახლოვის“ ალგორითმებმა, რომლებიც ემყარება სიტყვის დამახინჯების მარტივ ჰიპოთეზებს. ეს ნიშნავს, რომ მცდარი სიტყვა შეიცავს მხოლოდ ერთ შეცდომას, გაჩენილს ერთი ან ორი მეზობლად განლაგებული ლიტერის შეტანის მანუალურ პროცესში. მსგავსი დამახინჯების წყარო ასევე შესაძლოა იყოს სემანტიკური ან სინტაქსური შეთანხმების დარღვევის შემთხვევები. როგორც ე. წ. „სპელჩეკერების“ გამოყენების პრაქტიკა ცხადყოფს, მსგავსი შეცდომები ყველაზე ხშირია თითქმის ყველა ენისათვის, რაც, ბუნებრივია, არ გამორიცხავს სიტყვების დამახინჯების სხვა მოვლენებს და ეს მნიშვნელოვნად ართულებს აღნიშნულ მარტივ სქემას.

ამჟამად მსგავსი სისტემების კომერციულ ბაზარზე დე-ფაქტო სტანდარტად შეიძლება ჩაითვალოს სუგესტიურ ფორმათა შეთავაზების მითითებული პრინციპით შემუშავებული ისეთი ალგორითმები, როგორცაა „ერთი ლიტერის ჩასმა“- „ერთი ლიტერის წაშლა“- „ერთი ლიტერის შეცვლა“- „მეზობელი ლიტერების გადასმა“. „სპელჩეკერის“ შექმნისას მსგავსი სტანდარტული სუგესტიური ფორმების ეფექტურ საშუალებათა რეალიზაცია ერთ-ერთ ძირითად ამოცანას წარმოადგენს. საყურადღებოა, რომ ტექსტის დამახინჯების მოდელის პრინციპულად ევრისტიკული ხასიათის გამო ყველა სუგესტიურ წინადადებას უთუოდ უნდა დაემოწმოს ტექსტის ავტორი ან ოპერატორი კონტექსტის შესაბამისად, რათა საბოლოოდ მართებული ვარიანტი შეირჩეს.

ქართული მართლმწერის ინტერნეტვერსიისთვის ავტორების მიერ დამუშავებულია მონაცემთა შენახვის ქართულენოვანი სპეციალური სისტემა, რომელიც უზრუნველყოფს როგორც მართებული სიტყვაფორმის სწრაფ მოძიებას, ისე სუგესტიური ფორმების სწრაფ შეთავაზებას ტექსტის დამახინჯების სტანდარტული მოდელების ფარგლებში. შეთავაზებულ სისტემას საფუძვლად უდევს სიტყვების ნუსხის წარმოდგენა მრავალგანშტოებადი B+ ხის სახით, რომლის კვანძები შეიცავენ ქართული ანბანის ლიტერათა ნუსხას, რომლითაც იწყება უშუალოდ ამ ლიტერის შემდეგ სიტყვების დანარჩენი ნაწილი. თუ ლიტერის შესაძლო გამაგრძელებელი სიტყვის ფრაგმენტი ორი ან მეტია, ხის განმსაზღვრელი კავშირი ხორციელდება დაქვემდებარებული კვანძის საშუალებით. თუ შესაბამისი გაგრძელება ერთადერთია, – ის მოთავსდება უშუალოდ კვანძში ლიტერის სახით. ეს სტრუქტურა ემსგავსება ქართული ფუძეების ლექსიკური შემადგენლობის ანბანურად განლაგებას, სადაც მეზობლად მოხვედრილი სიტყვები, უმრავლეს შემთხვევაში, განსხვავდებიან მხოლოდ ბოლო ნაწილით.

ზუსტი ძეგლის შემთხვევაში მონაცემთა შენახვის მოცემული სტრუქტურისათვის ძირითადი საძიებო ოპერაციის (შედარებების, მეზობელი კვანძისკენ გადასვლების) ლოგარითმული შეფასება ნათელია. სუგესტიური ძეგლის შემთხვევაში, ტექსტის დამახინჯების მიღებული ჰიპოთეზიდან გამომდინარე, სიტყვის მართებული ფორმის ძეგლისათვის საჭირო განშტოება არ შეიძლება შესრულდეს იმაზე მეტჯერ, რამდენ ლიტერასაც შეიცავს დამახინჯებული სიტყვა. ეს ნიშნავს, რომ ხეზე გადაადგილებისას დაბრუნების რაოდენობა არ უნდა აღემატებოდეს სიტყვის სიგრძეს, ე.ი. ის არ არის დამოკიდებული შენახულ ფუძეთა რაოდენობაზე. აქედან გამომდინარე, ვინაიდან ქართული ენისათვის ფუძეთა საშუალო სიგრძე ცნობილია (არა უმეტეს 10-12 ლიტერისა) და შეიძლება ჩაითვალოს კონსტანტად, ძირითადი საძიებო ოპერაციის რაოდენობის ლოგარითმული შეფასება მისაღებია სუგესტიური ფორმების ძეგლის შემთხვევაშიც, თუმცა ეს რიცხვი საგრძნობლად იზრდება ზუსტი ძეგლის პროცესთან შედარებით.

კერძოდ, ქართული მართლმწერის ინტერნეტვერსიისათვის დამუშავებული ქართული სახელური ფუძეების მონაცემთა ბაზა შეიცავს 150,000-ზე მეტ ერთეულს, კვანძების რაოდენობა აღემატება 12,000-ს, კვანძის საშუალო სიგრძე მერყეობს 5-დან 10-მდე, თუ არ ჩავთვლით ხის ფუძესთან განლაგებულ რამდენიმე კვანძს. მნიშვნელოვნად უფრო კომპაქტურია თავისი რეგულარული სტრუქტურის გამო სახელური სიტყვაფორმების დაბოლოების მონაცემთა ბაზა, რომელიც შეიცავს 30,000 ერთეულს (ბრუნების სხვადასხვა ტიპისათვის).

სისტემის ალგორითმების განვითარებამ განაპირობა შემდეგი: მონაცემთა მართვის პრინციპის შეუცვლელად ტექსტის დამახინჯების ზოგიერთი რთული, “არამჭიდრო” ევრისტიკული მოდელისათვის ძეგლის ამოცანა გადაწყვეტდა ისე, რომ არ გაზრდილიყო ძეგლის ოპერაციის ლოგარითმული შეფასება. ამჟამად ქართული მართლმწერის ინტერნეტ-ვერსიაში რეალიზებულია შემდეგი ალგორითმები, რომლებიც ასახავენ ქართული სიტყვაწარმოების მრავალფეროვნების მნიშვნელოვან ასპექტებს:

- ორი მიმდებარე სიტყვის „შეწებების“ მოდელი (ამ სიტყვებს შორის შუალედის დაკარგვა), მაგ., *არაქვს* -> *არ აქვს*;
- ორი ლოგიკურად თანაბარი ფუძის გაერთიანების მოდელი სისტემის ბაზაში არარსებული კომპოზიტის სახით, რომლის კომპონენტებიც დეფისით გამოიყოფა, მაგ., *იანვარ-თებერვალი*;
- ორი ლოგიკურად დამოკიდებული ფუძის გაერთიანების მოდელი სისტემის ბაზაში არარსებული უდეფისო კომპოზიტის სახით, მაგ., *წერილმასშტაბიანი*;
- არსებითი სახელის წარმოების მოდელი სისტემისათვის ცნობილი თვისების განზოგადების საშუალებით, მაგ., *ბალაგანი* -> *ბალაგანურობა*;
- არსებითი სახელის წარმოების მოდელი სისტემისათვის ცნობილი თვისების მინიჭების საშუალებით, მაგ. *ბალაგანი* -> *ბალაგანური*.

The Means of effective search of Suggestive Forms in the internet version of the Georgian Spellchecker

Levan Chkhaidze, David Jashi

Technical University of Georgia
levan@dr.com

The search for the correct suggestive forms which would resemble the word to be determined is among the tasks which should be solved by the spellchecker. This resemblance is based on the errors made in spelling. In addition to the problem of the functioning of the spellcheck, this task seems to be problematic for automated spellcheck as well as due to the creation of other products of the programme for the complex systems such as the systems designed for search text databases using natural languages.

The search system algorithms for suggestive forms should be heuristic as they are based on various hypotheses taking into consideration the human character of the mistakes made. Among these are the algorithms of “close textual resemblance” based on the simple hypothesis of the word deviation. This means that the deviated word contains only one error arising in the process of inserting two neighbouring letters by hand. The similar distortion may also be caused by deviations in semantic or syntactical agreement. As revealed by the so called “spellcheckers” such mistakes belong to the most widely spread type. However,

this does not exclude other cases of deviation and thus turns the simple scheme into a complex one.

Currently, the algorithms based on the above mentioned principle of suggestive forms such as “insertion of one littera”, “deleting of one littera”, “replacement of one littera”, “inversion of the neighbouring litteras” can be considered to be a *de-facto* standard in the commercial market of similar systems. In the process of creating a “spellchecker” the realisation of an effective means of standard suggestive forms presents an important problem. On the other hand, because of the heuristic nature of the text distortion model all the suggestive sentences should be referred to the author of the text or the operator (according to the context) in order to select the correct variant.

The authors of this paper have elaborated a special system for storing the database of the internet version of the spellchecker which facilitates fast search of the correct word-form as well as a fast suggestion of suggestive forms within the frames of standard models of text distortion.

The proposed system is based on the presentation of the list of words as a ramificated tree “B+” the knots of which contain the list of litteras with the remaining words starting straight afterwards.

If there are two or more fragments of the possible word coming after the littera, the connection is made possible by means of the subordinate knot. If there is only one corresponding continuation it will be inserted in the knot as a littera. This structure resembles the alphabetic order of lexical constituents of the Georgian stem in which the neighbouring words differ mostly by the last part.

In the case of an exact search the logarithmic evaluation of the basic search operation (such as comparisons, movements towards the neighbouring knot) for the given structure of the saved data is clear. In the case of a suggestive search based on the hypothesis of the text distortion, the ramification necessary for the search of the correct form cannot be performed more often than the number of literas of the distorted word. This means that the number of backtracks to the tree should not be more than the length of the word which means that it does not depend on the number of the saved stems. Thus as the average length of stems for the Georgian language is known (not more than 10 - 12 litteras) and can be considered as constant the logarithmic estimation is acceptable in the case of a search for suggestive forms as well. However, this figure is much higher than that for the exact search. The database of Georgian nominal stems processed for the internet version of the spellchecker contains more than 150,000 units, the number of knots exceeds 12,000, the average length of the knot ranges between 5 - 10 not including several knots arranged at the stem of the tree.

The database of the endings of nominal word-forms is much more compact due to its regular structure. It includes 30,000 items (for different types of declension).

The development of algorithmic system made it possible to solve the problem of search for some “untight” complex heuristic models of text distortion without increasing the rates of logarithmic estimation of search operation.

Currently, the following algorithms depicting the crucial aspects of Georgian word-formation are realised in the internet version of the Georgian spellchecker:

- the model of “glueing” two neighbouring words (the loss of the interval between these words) *arakvas* → *ar akvs*;
- the model of uniting two logically equal stems realised by the composite lacking in the system the components of which are not separated by a hyphen: *ianvar-Tebervali*;
- the model of uniting two logically dependent stems realised in the system by the composite without the hyphen: *çvriļmasştabiani*;
- the model of formation of the noun by means of the generalization the property well-known to the system: *balagani* → *balaganuroba*
- the model of formation of the noun by means of conferring the property well-known to the system: *balagani* → *balaganuri*

ქართულის, როგორც უცხო ენის, დისტანციური სწავლების ძირითადი პრინციპები

ქეთევან გოჩიტაშვილი, გიული შაბაშვილი
თბილისის სახელმწიფო უნივერსიტეტი

kgochitashvili@yahoo.com

დისტანციური განათლება ერთ-ერთ პროგრესულ, თანამედროვე, იაფ, ხელმისაწვდომ და მოქნილ სისტემად ითვლება თანამედროვე საგანმანათლებლო სივრცეში. დასავლურ სამყაროში ამ ტიპის განათლების მიმართ ინტერესი ძალიან დიდია და მას უკვე ხანგრძლივი ისტორიაც აქვს.

არსებობს მრავალი კურსი სხვადასხვა ენის დისტანციურად შესასწავლად. თუმცა, უნდა აღინიშნოს, რომ ამ კურსების შექმნა და განხორციელება არ ხერხდება მხოლოდ კერძო ინიციატივით და მოითხოვს ინსტიტუციურ და, უფრო მეტიც, სახელმწიფო მხარდაჭერას, რადგან, გარდა კურსის შინაარსისა, რომლის შექმნაც ცალკეულ სპეცილისტს შეუძლია, მისი დიზაინი და ტექნიკური მხარდაჭერა, აგრეთვე მისი ხელმისაწვდომობის უზრუნველყოფა, მოითხოვს სოლიდური ფინანსური და ადამიანური რესურსების მობილიზებას.

რა ვითარებაა ამ თვალსაზრისით ქართულ საგანმანათლებლო სივრცეში და რა უნდა იქნეს გათვალისწინებული ქართულის, როგორც უცხო ენის, დისტანციური კურსის შექმნისას?

სამწუხაროდ, დღესდღეობით ქართულის, როგორც უცხო ენის, დისტანციური კურსების დეფიციტი აშკარაა და ამას თავისი მიზეზებიც აქვს. პოტენციური მომხმარებლის სიმცირე; არ არის სათანადო ინსტიტუციური მიდგომა და სახელმწიფო პოლიტიკა; თანამედროვე მეთოდოლოგიური ბაზის დეფიციტი.

რას უნდა აკმაყოფილებდეს ქართულის, როგორც უცხო ენის, სწავლების დისტანციური კურსი?

- კურსი დისტანციური სწავლების სტუდენტების მოთხოვნებს უნდა აკმაყოფილებდეს, უნდა ითვალისწინებდეს ისეთ ელემენტებს, როგორიცაა ონლაინ ლექციები, ლინკები დამატებითი ინფორმაციისათვის, კომუნიკაცია ელ. ფოსტის მეშვეობით კონსულტანტთან, დისკუსიები ჩეთის ან ფორუმების მეშვეობით, ტესტირება. კურსის ხანგრძლივობა 2 თვეზე მეტი არ უნდა იყოს. კურსი აგრეთვე უნდა მოიცავდეს კურსის შეფასების ფორმას, რომელიც სტუდენტების მიერ უნდა შეივსოს და შემდეგ გადაეგზავნოს დისტანციური სწავლების კოორდინატორსა და კურსის ხელმძღვანელს.

ქართულის, როგორც უცხო ენის, დისტანციური კურსის შექმნისას ავტორმა ან ავტორებმა უნდა გაითვალისწინონ შემდეგი ფაქტორები:

- კურსის ადაპტირება სოციოკულტურულ გარემოსთან;
- კურსის ადაპტირება ფსიქოლინგვისტურ გარემოსთან;
- კურსის დეტალური აღწერა;
- კურსის შექმნისა და სწავლების დროის განაწილების დეტალური აღწერა;
- კურსისა და კურსის მასალების გარკვეულობა, თითოეული გაკვეთილის მიზნების გათვალისწინებით;
- კურსის ორგანიზება და სტრუქტურა;
- დისტანციური სწავლების ელემენტების დართვა, როგორებიცაა ონლაინ ინტერაქტივობა მასწავლებელს და სტუდენტს შორის.

კურსის შინაარსი უნდა დაეყრდნოს ცნობილ საგანმანათლებლო მეთოდს – მარტივიდან რთულისკენ, თუმცა, სწავლების ტრადიციული, სწორხაზოვანი მეთოდისაგან განსხვავებით, დისტანციური, ონლაინ კურსი მოცემული უნდა იყოს სეგმენტირებულად, რათა იყოს ვიზუალურად ადვილად აღსაქმელი და იოლად მოსახმარი.

კურსის შინაარსი ორიენტირებული უნდა იყოს იმისკენ, რომ სტუდენტს განუვითაროს ენის ფლობის ოთხივე კომპეტენცია: წერა, კითხვა, მოსმენა, ლაპარაკი. ონლაინ რეჟიმში იძლევა იმის საშუალებას, რომ მსმენელი პარალელურად, სინქრონულად დაუფლოს ოთხივე კომპეტენციას, რაც კურსის შექმნელისაგან სათანადო კვალიფიკაციას, გამოცდილებასა და სწორი მეთოდოლოგიით მუშაობას მოითხოვს.

მოსმენა / ლაპარაკი – აუდიო და ვიდეორესურსები;

კითხვა / წერა სავარჯიშოვების სხვადასხვა ტიპები;

კურსის ფარგლებში ავტორებმა უნდა გაითვალისწინონ, რომ პირველ ეტაპზე აქტივობები მიმართონ რეპროდუცირებაზე, ხოლო პროდუცირება ეტაპობრივად დაგეგმონ. მნიშვნელოვანია იმის გათვალისწინებაც, რომ აღნიშნული პროცესი შესაძლოა მოცემული იქნეს ერთი თემის მსვლელობისას.

კურსის დაგეგმვისას უნდა გავითვალისწინოთ შემდეგი ეტაპები და პრინციპები:

- 1 **ეტაპი** - ენის ძირითადი ლექსიკური ფონდისა (აქტუალური ლექსიკა) და გრამატიკული სტრუქტურის შესწავლა (50 თემატური გაკვეთილი).
- 2 **ეტაპი** - ენის ლექსიკური ფონდის სწავლების გაღრმავება, სინონიმის, ფრაზეოლოგიისა და იდიომატიკის შესწავლა, ტექსტის (ქართულ კლასიკურ ნაწარმოებებზე დაყრდნობით) თავისუფალი კითხვა და გაგება (30 გაკვეთილი).
- 3 **ეტაპი** - ენის გრამატიკული სტრუქტურის ღრმად შესწავლა. ქართული ენის სინტაქსისა და სტილის საკითხების საკითხები. წერითი და მეტყველების კულტურის დახვეწა (50 გაკვეთილი).
- 4 **ეტაპი** - რიტორიკა და საქმიანი ქაღალდების ენა (30 გაკვეთილი).
- 5 **ეტაპი** - სამეცნიერო და აკადემიური სტილი (20 გაკვეთილი).

6 ეტაპი - ქართული ენის გრამატიკის თეორიული კურსი (ყველა გრამატიკული კატეგორიის მიმოხილვა, რომელიც ინტერნეტში **ლინკებით** ერთმანეთთან დაკავშირებული სტატიების სახით იქნება წარმოდგენილი);

დასასრულ, ყურადღებას გაავამახვილებთ ერთ-ერთ მნიშვნელოვან კომპონენტზე – დისტანციური სწავლებისას მასწავლებლის როლი მინიმუმამდე დაყვანილი, ის უნდა იყოს ტუტორი და არა ინსტრუქტორი, რომლის ფუნქციაც არის გააკონტროლოს სტუდენტის მხრიდან უკუგება, ამიტომ გაკვეთილის შინაარსი და დიზაინი ისე უნდა დაიგეგმოს და აიგოს, რომ სტუდენტს თვითშეფასების საშუალება ჰქონდეს (სწორია-არასწორია, ბონუსი, შემდეგ ეტაპზე გადასვლის უფლება და ა.შ.).

Main Principles of Distance Teaching of Georgian, as Foreign Language

Giuli Shabashvili, Ketevan Gochitashvili

Tbilisi State University

kgochitashvili@yahoo.com

Distance education is thought to be one of the most progressive, modern, cheap, available and flexible systems in contemporary educational space. The West is especially interested in this very system of education, so it has quite a long history.

Different languages possess many courses of distance learning. Creation and realization of these courses are beyond individual initiative, institutional and state support are demanded greatly. Indeed, one can compose a content, but design, technical support, availability and prominent financial and human resources as a support.

And what is about Georgian educational space, what items are to be considered while creation of distance course of teaching Georgian, as foreign language?

Unfortunately, deficiency of the distance courses of Georgian, as foreign language, is a reality. Several reasons dominate: scarcing of consumers, absence of proper institutional and state politics, modern methodology.

What should be the standarts for teaching of Georgian, as foreign language?

- ❖ demands of the students are to be satisfied. Several students can be included within. They are as follows: on-line lectures, links for additional information, consultation by e-mail, forum and chat discussins, testing. Duration of course depends. Courses are

strongly recommended to be evaluated by the students , they have to send back the special forms to coordinator of teaching and director of the course.

Several factors are to be taken into consideration by the teachers while composing distance teaching course for Georgian, as foreign language.

- ❖ adaptation of a course to socio-cultural environment;
- ❖ adaptation of a course to psycho-linguistic environment;
- ❖ detail description of a course;
- ❖ detailed time-management of teaching;
- ❖ exactness of a course and its material, taking into consideration the purposes of each lesson;
- ❖ structure of a course;
- ❖ on-line activities between a teacher and a student.

Classical method of education is to be applied, i. e. moving from simple conjunction to the complex one. Besides, on-line course must be presented divided into segments to make the course comfortable for visual perception.

Content of a course is destined to develop all four competences of knowledge of a language: writing, reading, listening, speaking.. Proper methodology would provide all four competences simultaneously.

listening/speaking - audio and video-resources.

reading/writing - different types of exercises.

Activities towards reproducing must for the first stage, next – that of producing. Certain topic can cover both activities.

The following stages and principles are to be taken into account while planning the course.

I stage _ teaching of Alphabet, basic lexical found and grammar structures of language.

II stage _ deepening of teaching into lexical founds, studying synonyms, phraseology and idioms, free reading and understanding of a text (depended on Georgian classical writings)

III stage _ profound study of grammar structure of language. Syntax and style of the Georgian language. Elaboration of culture of writing and speaking.

IV stage _ Rhetorics of business language. (elective)

V stage _ Scientific and academic style. (elective)

VI stage _ Theoretical course of Georgian grammar (survey of all grammar categories, linked up as articles).

Finally, it must be outlined that distance teaching diminishes the part of a teacher, ge/she is more a tutor, than instructor, controlling the student's comprehension.

სასტენდო მოხსენებები

**ქართული ენის და აზროვნების მათემატიკური თეორიის
სასაფუძვლო საკითხები**

და

**ხმით მართვადი ქართული ინტელექტუალური კომპიუტერული სისტემის
I ვერსიის ნაწილები**

კ. ფხაკაძე, გ. ჩიჩუა, ა. ვაშალომიძე, კ. გაბუნია, ლ. აბზიანიძე, ა. მასხარაშვილი, მ. ჩიქვინიძე
საქართველოს საპატრიარქოს წმინდა ანდრია პირველწოდებულის ქართული უნივერსიტეტი
ქართული ენის, ლოგიკისა და გამოთვლელის ღია ინსტიტუტი

მოხსენებისას მოკლედ მიმოვიხილავთ ქართული ენისა და აზროვნების მათემატიკური თეორიის შემუშავებისა და ხმით მართვადი ქართული ინტელექტუალური კომპიუტერული სისტემის კონსტრუირების მიზნით ჯგუფში წარმოებული თეორიული კვლევების ძირითად შედეგებს. კერძოდ, ყურადღება გამახვილდება:

1. იმ ახალ ენობრივი იდეოლოგიაზე, რომელსაც ჩვენ ქართულ ენობრივ იდეოლოგიას ვუწოდებთ;
2. ქართული ენის ზოგად და კერძო თეზისებზე;
3. ქართული ენისა და აზროვნების მათემატიკური თეორიის სასაფუძვლო საკითხებზე;
4. ქართული მეტყველების მასინთეზებელი და გამომცნობი სისტემების შემუშავების მიზნით ჯგუფში წარმოებული თეორიული კვლევების შედეგებზე.

გარდა ამისა, მოხსენებისას წარმოვადგენთ ხმით მართვადი ქართული ინტელექტუალური კომპიუტერული სისტემის იმ ქვესისტემებს, რომელთა წინასწარი ვერსიები უკვე შემუშავებულია. ესენია:

1. ქართული გრამატიკული მართლმწერი სისტემა;
2. ქართულ-მათემატიკური მთარგმნელი სისტემა;
3. ქართული მრავალენოვანი ცალმხრივ მთარგმნელი სისტემა;
4. ქართული მკითხველი სისტემა;
5. ხმისჩამღები საშუალებებიანი ქართული მკითხველი სისტემა;
6. ლექსიკურად შეზღუდული ქართული მეტყველების ამომცნობი სისტემა;
7. ქართული ხმით მართვადი ხმოვანი ინტერნეტი (მოკლედ - ქართული ხმოვანი ინტერნეტი).

საზოგადოდ, ჩვენი კვლევები მიმართულია ქართული ენისა და აზროვნების სრული მექანიკური (ე.ი. კომპიუტერული) და მათემატიკური (ე.ი. გამოთვლითი) მოდელების შემუშავებისკენ. კერძოდ, ჩვენი კვლევები მიეკუთვნება მათემატიკის დაფუძნებასა და მათემატიკურ ლოგიკას – უფრო მეტად ბუნებრივი ენების ლოგიკას, ლინგვისტიკას – უფრო

მეტად გამოთვლით და მათემატიკურ ლინგვისტიკას, კომპიუტერულ მეცნიერებებს – უფრო მეტად ხელოვნური ინტელექტის ამოცანას. უფრო დეტალურად:

1. ჩვენ ფრეგესეული, მონტეგეიუსეული და ჩომსკისეული ზოგადი მიდგომების გათვალისწინებით, ვიკვლევთ ბუნებრივი ქართული ენის ლოგიკას და, ამავდროულად, ვცდილობთ ქართული ენისთვის მათემატიკური მიდგომებით აგებული თანამედროვე ლოგიკური გრამატიკის შემუშავებას;

2. ჩვენ ვიკვლევთ ქართულ წერით, ზეპირ და აზრით ენებს როგორც ცალ-ცალკე, ასევე ერთმთლიანობაში, ანუ მათი განუყოფელი ურთიერთკავშირებისა და მჭიდრო ურთიერთმიმართებების გათვალისწინებით და, ამავდროულად, ვცდილობთ ქართული ენისა და აზროვნების მათემატიკური თეორიის შემუშავებას: ანუ, სხვა სიტყვებით რომ ვთქვათ, ჩვენ ვცდილობთ ქართული ენობრივი აზროვნების იზომორფული მათემატიკური ტიპის კვაზი-ფორმალური ინტელექტუალური თეორიის აგებას;

3. ჩვენ ვიკვლევთ ქართული ზეპირი ენის ბუნებრივ თავისებურებებს და, ამავდროულად, ვცდილობთ ქართული მეტყველების მასინთეზებელი და ამომცნობი სისტემების კონსტრუირებას;

4. ჩვენი მიზანია კომპიუტერის აღჭურვა ქართული მეტყველებითი, სმენითი და აზროვნებითი უნარებით: ანუ, ჩვენი მიზანია ქართული ენობრივი აზროვნების იზომორფული კვაზი-ფორმალური ინტელექტუალური მათემატიკური თეორიითა და ქართული მეტყველების მასინთეზებელი და გამომცნობი სისტემებით კომპიუტერის სრული პროგრამული უზრუნველყოფა.

ეს კვლევები დაიწყო 1999 წელს და 2003 წლიდან მოყოლებული ის მიმდინარეობს თსუ სახელმწიფო მიზნობრივი პროგრამით „კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“ განსაზღვრული და ზემოთ მოკლედ გაცხადებული მიზნების გათვალისწინებით. პროექტი „კომპიუტერის სრული პროგრამულ-მომსახურებითი მოქცევა ბუნებრივ ქართულ ენობრივ გარემოში“, რომლის ძირითადი მიზანი, როგორც უკვე აღინიშნა, არის ქართული ენის აღჭურვა თანამედროვე და პერსპექტიული ტექნოლოგიური შესაძლებლობებით, სახელმწიფო მიზნობრივ პროგრამად 2002 წელს განისაზღვრა. ამჯერად ჩვენ ვამუშავებთ პროექტს „ქართული ენის და აზროვნების მათემატიკური თეორიის სასაფუძვლო საკითხები და ხმით მართვადი ქართული ინტელექტუალური კომპიუტერული სისტემის I ვერსიის ნაწილები“, რომელსაც საქართველოს საპატრიარქოს წმინდა ანდრია პირველწოდებულის ქართული უნივერსიტეტი აფინანსებს და რომელიც ზემოაღნიშნული სახელმწიფო მიზნობრივი პროგრამის ამ ეტაპობრივი ქვეპროგრამაა.

ქვემოთ მოკლედ წარმოვადგენთ ჩვენი საკონფერენციო მოხსენების ძირითად თეორიულ საკითხებს:

1. ჩვენი კვლევების ერთ-ერთი უმნიშვნელოვანესი შედეგია ის ენობრივი იდეოლოგია, რომელსაც ქართული ენობრივი იდეოლოგია ვუწოდებთ: ამ იდეოლოგიას ეწოდა ქართული ენობრივი იდეოლოგია იმის გამო, რომ ბუნებრივი ქართული ენის თავისთავადი მათემატიკური ბუნება საკმარის საფუძველს იძლევა ამ ხედვების მართებულობაში დასარწმუნებლად: ქართული

ენობრივი იდეოლოგიის საფუძველზე ჩვენ ვაცხადებთ **(i) დებულებას** და მას ჩვენ **ქართული ენის ზოგად თეზის** ვუწოდებთ:

(i) ნებისმიერი ბუნებრივი ენა და ამ ენით მოცემული აზროვნება არის პირველადი მათემატიკური თეორიის (პმთ) ბუნებრივი გაფართოების შედეგი: აქ პმთ არის კვაზი-ფორმალურად გაფართოებადი, ანუ ენობრივად განვითარებადი ევკლიდეს ტიპის კვაზი-ფორმალური მათემატიკური თეორია, რომლის ძირითადი ალფაბეტი პირველად მათემატიკურ ენად (პმე) იწოდება. ამასთან, ჩვენი ხედვებით, ამ თეორიის ძირითადი ცნებები, აქსიომები და გამოყვანისა და გაფართოების ზოგადი წესები ადამიანებში უნივერსალურად არსებული პირველადი შინაარსობრივი მონაცემებია, რის გამოც მათ ჩვენ პირველად მათემატიკურ კონცეპტებს (პმკ), ანუ თანდაყოლილ სააზროვნო მონაცემებს (თსმ) ვუწოდებთ.

2. კიდევ ერთი მნიშვნელოვანი შედეგი ჩვენი თეორიული კვლევებისა, რომლებიც ქართული ენის ბუნებრივი თავისებურებების გარდა ძირეულად ეყრდნობა აგრეთვე პროფ. შ. ფხაკაძისეულ აღნიშვნათა თეორიას, მისეულ ზოგად სემანტიკურ პროგრამასა და საკმარისად ზოგად **Ⅴ** მათემატიკურ ენას, არის **(ii)** და **(iii) დებულებები**, რომლებიც **ქართული ენის კერძო თეზისების** სახით უკვე გაცხადდა 2008 წელს ივ. ჯავახიშვილის სახელობის თბილისის სახელმწიფო უნივერსიტეტის 90 წლისა და ი. ვეკუას სახელობის გამოყენებითი მათემატიკის ინსტიტუტის 40 წლის საიუბილეოდ მიძღვნილ საერთაშორისო კონფერენციაზე „**გამოყენებითი მათემატიკის თანამედროვე პრობლემები**“:

(ii) ქართული ენა არის პირველადი მათემატიკური ენის ფორმალური გაფართოების შედეგი.

(iii) ქართული ენობრივი თეორია არის საკმარისად ზოგადი **Ⅴ მათემატიკურ თეორიის ქართული მოდიფიკაცია.**

3. მოხსენებისას მიმოვიხილავთ წერითი, ზეპირი და აზრითი ენების ურთიერთკავშირებს და მოკლედ წარმოვადგენთ ქართული ენისა და აზროვნების ძირეულ ნაწილს როგორც კვაზი-ფორმალურ **Ⅴ** მათემატიკურ თეორიას. ასევე, მოკლედ წარმოვადგენთ ამ თეორიის ძირითად ცნებებს, აქსიომებს და გამოყვანისა და გაფართოების წესებს. გარდა ამისა, საკმარისი სისრულით დასაბუთდება, რომ ქართული წერითი ენა არის ქართული აზრითი ენის ფორმალური გაფართოების შედეგი. ამასთან, ყურადღება გამახვილდება ქართული ენისა და აზროვნების მათემატიკური თეორიის ისეთ სასაფუძვლო საკითხებზე როგორებიცაა ქართული თხრობითი კილოს ზმნების, ზმნიზელების, არსებითი, ზედსართავი და განმაზოგადებელი სახელების, ბრუნების, უღლების, გრამატიკული მორფემების, სინტაქსური და კვაზისინტაქსური წყვილებისა და მარტივი და რთული წინადადებების მათემატიკურ გაგებებსა და განსაზღვრებებზე.

4. მოხსენებისას მოკლედ გაშუქდება კიდევ ერთი ხაზი ჩვენი კვლევებისა, რომლებიც ქართული მეტყველების მასინთეზებელი და გამომცნობი სისტემების კონსტრუირების მიზნით წარმოებს: ჯგუფში ორი განსხვავებული მიდგომით უკვე აგებულია ორი განსხვავებული სახის მკითხველი სისტემა. აგრეთვე, ჯგუფში უკვე აგებულია ხმისჩამდები საშუალებებიანი მკითხველი სისტემის ექსპერიმენტული ვარიანტი. ასევე, ჯგუფში აგებულია ლექსიკურად შეზღუდული ქართული მეტყველების გამომცნობი სისტემა. გარდა ამისა, მოხსენებისას

მიმოვიხილავთ ხმისჩამღებ საშუალებებიან ქართულ მკითხველ-მსმენელ სისტემას, რომლის აგებაზე ჯგუფში უკვე მიმდინარეობს მუშაობა. ეს სისტემა, რომლის ერთ-ერთი ძირითადი ქვესისტემა არის ხმისჩამღები საშუალებებიანი ქართული მკითხველი სისტემა, მეტყველების გამოცნობისას ისარგებლებს შინაგანი გასმენის პრინციპით. ეს გულისხმობს, რომ სისტემა მასში უკვე ჩადებული მომხმარებლის ხმის მეშვეობით შეძლებს ამ მომხმარებლის მიერ გამეტყველებული ფრაზის შინაგან გენერირებას, ანუ გასმენას და ამის საფუძველზე მის გამოცნობას. – მოხსენებისას ამ საკითხთან დაკავშირებით, როგორც დასაწყისშივე აღინიშნა, წარმოდგენილი იქნება ქართული ხმოვანი ინტერნეტი, რომლის ძირითადი ნაწილებია ზემოთ მოკლედ მიმოხილული ქართული მეტყველების მასინთეზებელი და გამომცნობი სისტემები.

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

**Pkhakadze K. Sh., Chichua G. T., Vashalomidze A. G., Gabunia K. A.,
Abzianidze L. V., Maskharashvili A. K., Chikvinidze M. L.**

St. Andrew the First Called Georgian University of The Patriarchy of Georgia
The Open Institute of the Georgian Language, Logic, and Computer

At the presentation we will overview the main results of our researches, which are currents for the aims of elaboration of the mathematical theory of the Georgian language and thinking and construction the talkie Georgian intellectual computer system. More strictly, at the presentation we will concentrate:

1. On the lingual ideology, which we call as Georgian Lingual Ideology;
2. On the Georgian language's theses;
3. On the some grounding questions of the mathematical theory of the Georgian language and thinking;
4. On the results of that our researches, which are carried on for the aims to create the speech synthesizing and recognizing technology for the Georgian language.

Also, at the presentation we will demonstrate some those subsystems of the 1st Version of the Voice Managed Georgian Intellectual Computer System, pre-versions of which we have already constructed. These are:

3. 1. Georgian Grammatical Spellchecker System;
3. 2. Georgian-Mathematical Translator System;

3. 3. Georgian Many-Lingual one way Translator System;
3. 4. Georgian Text Reader System with Fixed Voice;
3. 5. Georgian Text Reader System with User's Possibility to Build in an Own Synthetic Voice;
3. 6. Georgian Restricted Speech Recognition System;
3. 7. Georgian Voice Managed Internet with Voice, i.e. Georgian Talkie Internet;

Generally, our researches aim for the complete mechanical (i.e. computer) and mathematical (i.e. computational) modeling of the Georgian language and thinking. More strictly, our researches belong to the foundations of mathematics and mathematical logic - mainly to the logic of natural languages, to the linguistics – mainly to the computational and mathematical linguistics, and to the computer science – mainly to the artificial intelligence. In more details:

1. We are studying the natural logic of the Georgian natural language and we are trying to elaborate the contemporary mathematical type logical grammar for the Georgian language on the basis of Frege's, Montague's and Chomsky's general approaches;

2. We are studying natural Georgian written, spoken and thinking languages as differently, as well as in their inter-relation and we are trying to elaborate the mathematical theory of the natural Georgian language and thinking - in other words, we are elaborating the mathematical type quasi-formal intellectual theory, which will be isomorphic to the natural Georgian lingual thinking;

3. We are studying the natural Georgian spoken language for the aim to create the modern speech synthesizer and recognizer systems for the natural Georgian spoken language;

4. We are aiming to equip a computer with the natural Georgian speaking, listening and thinking abilities through the softwearing it with the mathematical type quasi-formal intellectual theory of the natural Georgian lingual thinking and with the Georgian speech synthesizer and recognizer systems.

These researches are begun in 1999 and since 2003 they go under the theoretical and technological aims of the State Priority Program (SPP) "Free and Complete Inclusion of a Computer in the Georgian Natural Language System", which aims are shortly declared above.

The program "Free and Complete Inclusion of a Computer in the Georgian Natural Language System", which is aimed on developing a modern technological tools for the Georgian language, was proved as SPP in 2002.

By now we are working out the project "Foundations of Mathematical Theory of the Georgian Language and Thinking and Sub-systems of the 1st Version of the Voice Managed Georgian Intellectual Computer System", which is a current subproject of the SPP "Free and Complete Inclusion of a Computer in the Georgian Natural Language System" and which is funded by the St. Andrew the First Called Georgian University of The Patriarchy of Georgia.

Below, we will shortly describe main theoretical topics of our presentation:

1. A very important result of our researches is that new lingual ideology, which we call as a Georgian Lingual Ideology (GLI). - It must be underlined: We call this lingual ideology as GLI because of the natural mathematical specifics of the Georgian language give us sufficient arguments to believe in it. – On basis of the GLI we declare (i) as Georgian language's general thesis:

(i) Any natural language and thinking, shortly any natural lingual thinking is a result of natural extension of Primary Mathematical Theory (PMT): Here, the PMT is a quasi-

formally extendable (i.e. symbolically developable) quasi-formal Euclid type axiomatic theory, which language (basic alphabet) is called as Primary Mathematical Language (PML) and which basic notions, axioms, general rules of inference, and general rules of extension are in all humans universally existed primary semantic dates, because of which they are called as Primary Mathematical Concepts (PMC), or as Inborn Semantic Data (ISD).

2. The very important results of our theoretical researches, which besides the natural Georgian lingual specifics are based on Prof. Sh. Pkhakadze's notation theory, on his general semantic program, and, also, on his sufficiently general \mathfrak{S} mathematical language, are **(i)** and **(ii)**, which we have already declared as Georgian language's theses in 2008 at the international conference on **"Modern Problems in Applied Mathematics"** dedicated to the 90th anniversary of Iv. Javakhishvili Tbilisi State University and 40-th anniversary of I. Vekua Institute of Applied Mathematics: It must be underlined that differently from above declared Georgian language's general thesis we have already sufficiently proved Georgian language's theses, which are state below:

(i) Georgian Language is a result of formally extension of the PML.

(ii) Georgian Lingual Theory is a Georgian type modification of the sufficiently general \mathfrak{S} mathematical theory.

3. At the presentation we will shortly overview inter-relation between Georgian written, spoken and thinking languages and there will be described the Core Part (CP) of the Georgian language and thinking as a quasi-formal \mathfrak{S} mathematical theory. Also, we will present some basic notions, some axioms, some rules of inference and some rules of extensions of this \mathfrak{S} mathematical theory. In addition, there will be shown that the CP of the Georgian Written Language (GWL) is a formal extension of the CP of Georgian Thinking Language (GTL). Besides above mentioned, at the presentation we will concentrate on such grounding questions of the mathematical theory of the Georgian language and thinking which are mathematical definitions of the Georgian declarative verbs, nouns, adjectives, adverbs, quantifiers, Georgian declination system, Georgian conjugation system, Georgian syntax and quasi-syntax pairs, Georgian simple and complex sentences, and, also, we will concentrate on mathematical nature, i.e. on mathematical definitions of different type Georgian grammatical morphemes.

4. The another trend of our researches is performed for the aims to create the speech synthesizing and recognizing technology for Georgian spoken language. By now, with the help of different approaches we have already constructed two different Georgian text reader systems with fixed voice and the first experimental version of the Georgian text reader system with user's possibility to build in an own synthetic voice. Also, we have already constructed Georgian speech recognition system, which recognizes Georgian speech with 97 percent of recognition in restricted lexicon with 300 words. In addition of above mentioned, by now we are elaborating the Georgian reader-listener system with user's possibility to build in an own synthetic voice. Here, it must be underlined, that the Georgian text reader system with user's possibility to build in an own synthetic voice is a one of the main sub-system of the Georgian reader-listener system with user's possibility to build in an own synthetic voice: We try to solve the speech recognition problem for Georgian basing on the method of internal listening. According to this method for the recognition

of a Georgian word pronounced by some Georgian speaker the Georgian reader-listener system with user's possibility to build in an own synthetic voice will use its ability to synthesize the natural spoken form of this Georgian word in the natural spoken voice of this Georgian speaker. - To this direction, as was mentioned, at the presentation we will present above underlined Georgian Talkie Internet: This will help us to make short presentation of the Georgian Text Reader System with Fixed Voice, of the Georgian Text Reader System with tools to Build in it Users Synthetic Voices, of the Georgian Speech Recognition System, which, and it is a clear, are main subsystems of the Georgian Talkie Internet.

ბესარიონ სოლოლაშვილი, ზაზა ინჯია

ინტერნეტის ქართულ ენაზე გამახმოვანებელი პროგრამა "ინტერბუ".

მხედველობის უქონლობა მძიმე დაბრკოლებას წარმოადგენს უსინათლო ადამიანებისათვის მათი შრომითი დასაქმებისა და ინტელექტუალური ზრდისათვის. ავტორების მთავარ მიზანს, აგერ უკვე შვიდი წელიწადია, შეადგენს შექმნას ისეთი გარემო-პირობები რომელიც ნაწილობრივ მაინც გამოასწორებს ამ ხარვეზს და თავიდან აგვაცილებს უსინათლოების მარგინალიზაციას. საშუალებას მისცემს მათ დამოუკიდებლად შეძლონ კითხვა, წერა და ინფორმაციის მოძიება კომპიუტერის საშუალებით.

2004-2007 წლებში შევქმენით ქართული ტექსტების გამახმოვანებელი პროგრამა „ბუ“, რომელიც საშუალებას იძლევა უსინათლოებმა მოისმინონ წიგნები, დამოუკიდებლად აკრიფონ ტექსტობრივი მასალა, შეასწორონ და ამოებუდონ პრინტერზე. ამ პროგრამამ საშუალება მოგვცა საქართველოს რამდენიმე მთავარ ბიბლიოთეკაში გახსნილიყო უსინათლოებისათვის სამკითხველო დარბაზები.

თანამედროვე ეტაპზე ინფორმაცია მიიღება არა მარტო წიგნების საშუალებით, არამედ სულ უფრო და უფრო აქტუალური ხდება ინფორმაციის ინტერნეტით მოპოვება და კომუნიკაცია, რამაც დღის წესრიგში დააყენა უსინათლოთათვის ადაპტირებული ინტერნეტით სარგებლობის პროგრამის შექმნის აუცილებლობა. ქართული ინტერნეტის გამახმოვანებელი პროგრამა „ინტერბუ“ პროგრამა „ბუსთან“ ერთად შექმნის უსინათლოებისათვის კომპიუტერის გამოყენების მეტ-ნაკლებად სრულ პაკეტს.

პროგრამა "ინტერბუ" შეიცავს შემდეგ ქვესისტემებს:

1. საძიებო ქვესისტემას;
2. ვებ-გვერდების გამოსაძახებელ და გამახმოვანებელ ქვესისტემას;
3. ელ-ფოსტის გამოყენების ქვესისტემას;

4. სკაიპს.

პროგრამის ვებ გვერდების მიღება ხორციელდება ვინდოვს ინტერნეტ ექსპლორერის კომპონენტის გამოყენებით. ხმოვანდება როგორც ვიზუალური ვებ გვერდი (ძირითადად ბმულებზე ნავიგაციით) ასევე მისი სრული ტექსტური ნაწილი ტექსტური ნავიგაციის და რედაქტირების საშუალებით. ხმოვანდება ინტერაქტიული რეჟიმი (ვებ გვერდზე ინფორმაციის შეტანისას, მაგალითად - პაროლი) თუმცა გვხვდება ხოლმე პრობლემები ამ რეჟიმებში კლავიატურით მუშაობისას, იმ მიზეზით, რომ ვებ გვერდის შემქმნელი ზრუნავს მის ვიზუალურ მხარეზე და ელემენტები შეიძლება მოათავსოს შიდა სტრუქტურაში არათანმიმდევრულად. აგრეთვე ვერ ხერხდება გრაფიკული ტექსტის გახმოვანება და რთულდება ქვეპროგრამებით მიღებული ზოგი ვიზუალური ელემენტის გახმოვანება.

პროგრამით შესაძლებელია რჩეული ვებ გვერდების მისამართების დაგროვება, გვერდზე არსებული ბმულების სიის მიღება, ძიება გვერდზე და ინტერნეტში (ვიყენებთ გუგლის საძიებო სისტემას).

პროგრამა შეიცავს ელექტრონული ფოსტის ძირითად ელემენტებს: წერილების და ფაილების გაგზავნა-მიღება, მისამართების სია, მომხმარებლის სხვადასხვა მეილ-სერვერით სარგებლობის საშუალება.

სკაიპით სარგებლობისას მომხმარებელს შეუძლია მოიძიოს მისამართი, დარეკოს ან აწარმოოს გახმოვანებული მიწერ მოწერა.

ქართულის გარდა გახმოვანება ხორციელდება ინგლისურ და რუსულ ენაზე (ამ ენებისთვის კომპიუტერში ჩატვირთული უნდა იყოს შესაბამისი გახმოვანების ძრავები).

Besarion Sologashvili, Zaza Injia

Internet text-to-speech Software INTERBU

Blindness is very hard barrier for people for working and intellectual growing.

Main aim of authors is to create helper tools for blind to partially solve their problems and to allow them writing, reading and looking for information using computer.

In 2004-2007 we have designed Georgian text-to-speech software BU. This program allows blind users to listen books, to type texts, to correct and print. This program is used by blinds in main libraries of Georgia.

Today besides other ways the people takes information and communicates with Internet. Therefore adaptation of internet possibilities for blinds is very actual problem. Georgian text-to-speech software with internet solutions can be used as more completed package.

The program INTERBU includes these units:

1. Searching system;
2. Web page speech-browsing system;
3. E-mail system;
4. Skype system;

Accessing to web pages succeed by Windows Internet Explorer Activex.

Collecting of favorite web pages, links on web pages and searching is possible. We use Google search engine.

Program includes main elements of mailing: sending and receiving mails and files, address book and user accounts organizer.

User can find addresses by Skype, make call and chat by speech.

Besides Georgian language user can switch voice in English or Russian speech (Microsoft speech engine must be installed for these languages).