

**სასრული მდგომარეობის ავტომატის კომპაილერი
ქართული ენის მორფოლოგიური პროცესორისათვის**

ლიანა ლორთქიფანიძე

liana.lortkipanidze@tsu.ge

კომპიუტერული მეცნიერებების დეპარტამენტი, ივანე ჯავახიშვილის სახელობის თბილისის
სახელმწიფო უნივერსიტეტი, უნივერსიტეტის 13

არ არსებობს ბუნებრივი ენების სრული ლექსიკონები. აგრეთვე შეუძლებელია გადაითვალოს რიცხვების უსასრულო სიმრავლე, ან დასახელდეს ყველა არსებული საკუთარი სახელი და გვარი. დროთა განმავლობაში ყველა ენა იცვლება. იცვლება მისი ლექსიკონიც. ენის ყოველ ქვესისტემას თავისი საკუთარი გამოხატვა აქვს და შეუძლებელია ენის ლექსიკონში შევიდეს მისი ყველა ცალკეული ქვესისტემის სრული ლექსიკა. ქართული ენა 17-მდე დიალექტითაა წარმოდგენილი. მიმდინარეობს მუშაობა ქართული ენის დიალექტების კორპუსზე. კორპუსის სამუშაოთა ტექნოლოგიური ჯაჭვის ერთი ნაწილია მოხსენებაში დასმული საკითხი.

ამჯერად ყურადღება გვინდა გავამახვილოთ ენის დიალექტური მორფოლოგიური პროცესორის კომპილირებაზე.

ჩვენ წარმოგიდგენთ საკმაოდ მარტივ და, ამავე დროს, სრულყოფილ ტექნიკას, რომელიც სალიტერატურო ენის უკვე არსებული მორფოლოგიური პროცესორის ადაპტაციით იძლევა დიალექტური ენის პროცესორის კომპილირების საშუალებას. ჩვენი კომპიუტერული პროგრამის ინსტრუმენტების (ხელსაწყოების) დახმარებით ხდება ენის სხვადასხვა დიალექტისთვის სისტემის გაწრთვნა (ტრენინგი) ცნობილი მორფო-ფონემური წესების გამოყენებით.

ჩვენი მეთოდის შესამოწმებლად შევარჩიეთ ქართული ენის დიალექტების კორპუსი. ქართული სალიტერატურო ენის მორფოლოგიურ მოდელზე დაყრდნობით მოვახდინეთ მორფოლოგიური პროცესორის ადაპტაცია და შემდეგ მოვსინჯეთ დიალექტური კორპუსის ლემატიზაცია და ზედაპირული ანოტირება.

ნაშრომის მე-2-ე პარაგრაფში განხილული იქნება ენის მორფოლოგიური პროცესორის კომპილირების სისტემა ქართული ენის მაგალითზე, მე-3-ე პარაგრაფში განვიხილავთ ლიტერატურული ენის პროცესორის ადაპტაციის ტექნიკას დიალექტური ქვესისტემებისთვის, მე-4-ე პარაგრაფში აღწერილი იქნება მორფოლოგიური ანალიზატორი და მე-5-ე პარაგრაფში შევეხებით მსგავს სამუშაოებს.