

ბათუმის შოთა რუსთაველის სახელმწიფო  
უნივერსიტეტი  
ჰუმანიტარულ მეცნიერებათა ფაკულტეტი  
ევროპეისტიკის დეპარტამენტი

ხელნაწერის უფლებით

ხათუნა დიასამიძე

იურიდიული ტექსტების თარგმანის ანალიზის  
მოდელირება პარალელურ კორპუსში

სპეციალობა - თარგმანმცოდნეობა

ფილოლოგიის დოქტორის აკადემიური ხარისხის  
მოსაპოვებლად წარმოდგენილი  
სადისერტაციო ნაშრომის

ა ნ ო ტ ა ც ი ა

ბათუმი - 2021

სადისერტაციო ნაშრომი შესრულებულია ბათუმის შოთა რუსთაველის სახელმწიფო უნივერსიტეტის ჰუმანიტარულ მეცნიერებათა ფაკულტეტის ევროპეისტიკის დეპარტამენტში.

**სამეცნიერო ხელმძღვანელი: ხათუნა ზერიძე**

ბათუმის შოთა რუსთაველის სახელმწიფო უნივერსიტეტის პროფესორი, ფილოლოგიის დოქტორი.

**შემფასებლები:**

1. **თინათინ მარგალიტაძე** - ფილოლოგიის მეცნიერებათა დოქტორი, ილიას სახელმწიფო უნივერსიტეტის პროფესორი;
2. **თინათინ ბოლქვაძე** - ფილოლოგიის მეცნიერებათა დოქტორი, თსუ-ს თეორიული და გამოყენებითი ენათმეცნიერების სასწავლო-სამეცნიერო ინსტიტუტის ასოცირებული პროფესორი;
3. **ლია ქაროსანიძე** - ფილოლოგიის მეცნიერებათა დოქტორი, თსუ-ს არნ. ჩიქობავას სახელობის ენათმეცნიერების ინსტიტუტის სამეცნიერო ტერმინოლოგიისა და თარგმნითი ლექსიკონების განყოფილების ხელმძღვანელი.

**საერთაშორისო ექსპერტი:**

**ეკატერინა ლაფზინოვა-კოლტუნსკი** - ფილოსოფიის დოქტორი, საარლენდის უნივერსიტეტის პროფესორი (საარბრუკენი, გერმანია).

სადისერტაციო ნაშრომის დაცვა შედგება 2021 წლის 26 ივლისს 17 საათზე, N37 აუდიტორიაში, ბათუმის შოთა რუსთაველის სახელმწიფო უნივერსიტეტის ჰუმანიტარულ მეცნიერებათა ფაკულტეტის სადისერტაციო საბჭოს მიერ შექმნილი სადისერტაციო კომისიის სხდომაზე.

მისამართი: ბათუმი 6010, ნინოშვილის/რუსთაველის ქ. 35/32. სადისერტაციო ნაშრომის გაცნობა შეიძლება ბათუმის შოთა რუსთაველის სახელმწიფო უნივერსიტეტის ილია ჭავჭავაძის ბიბლიოთეკაში, ხოლო სადისერტაციო ნაშრომის ანოტაციისა - ამავე უნივერსიტეტის ვებ-გვერდზე ([www.bsu.edu.ge](http://www.bsu.edu.ge)).

**სადისერტაციო საბჭოს მდივანი,**  
ფილოლოგიის დოქტორი, ასოცირებული პროფესორი **მ. კიკვაძე**

## შესავალი

კორპუსული კვლევა თანამედროვე თარგმანმცოდნეობაში დღესდღეობით საკმაოდ აქტუალური გახდა. თარგმანმცოდნეობისა და კორპუსის ურთიერთკავშირი ჯერ კიდევ გასული საუკუნიდან იღებს სათავეს. წლების განმავლობაში კომპიუტერული ტექნოლოგიების განვითარებასთან ერთად კორპუსზე დაფუძნებული თარგმანმცოდნეობა გაფართოვდა და თანდათანობით განსხვავებული სფეროები მოიცვა, ჩამოყალიბდა სხვადასხვა სახის პარალელური კორპუსი. საკუთრივ „პარალელური კორპუსი გულისხმობს ორ ან რამდენიმე ენაზე არსებული სრულიად იგივეობრივი ტექსტების მოთავსებას სპეციალურად შემუშავებულ კომპიუტერულ პროგრამაში და ურთიერთშედარებისა და ანალიზის საფუძველზე მათ შორის არსებული ეკვივალენტობის დადგენას“ (Yang, 2002:29). პარალელურ კორპუსს განსაკუთრებული მნიშვნელობა ენიჭება თარგმანმცოდნეობისათვის, ვინაიდან ის აკავშირებს წყარო და სამიზნე ენებს როგორც ტექსტების, ასევე წინადადებებისა და სიტყვების დონეზე, თანადროულად წარმოაჩენს მათ მთარგმნელის წინაშე და აადვილებს მათს საქმიანობას. მეორე მხრივ, როგორც მოგეხსენებათ, არსებობს ეროვნული კორპუსები, რომლებიც კონკრეტული მიზნებით შეიქმნა, ასეთებია, მაგალითად, ინგლისური ენის ან ქართული ენის ეროვნული კორპუსები. ისინი აურაცხელ ენობრივ მასალას შეიცავენ და დიდ როლს ასრულებენ ამა თუ იმ ტექსტის ლინგვისტური თუ სტილური თავისებურებების განსაზღვრაში. არსებობს მთარგმნელობითი კორპუსული ლექსიკონები, მაგ. პროფესორ თ. მარგალიტაძის თაოსნობით თსუ ლექსიკოგრაფიის ცენტრში შექმნილი დიდი ინგლისურ-ქართული ლექსიკონი [www.dict.ge](http://www.dict.ge) სწორედ განმარტებების კორ-

პუსული სისტემატიზაციის პრინციპით არის შექმნილი. არსებობს ასევე მთელი რიგი დარგობრივი (სამედიცინო, იურიდიული, პოლიტიკური და სხვა) კორპუსები. თანდათანობით გამოიკვეთა კვლევის საერთო ობიექტები, რამაც განაპირობა ზემოაღნიშნული კორპუსების საფუძველზე პარალელური კორპუსების შექმნა, რაც განსაკუთრებით მნიშვნელოვანია თარგმანმცოდნეობისთვის. იურიდიული კორპუსები დღესდღეობით არ არის საბოლოოდ ჩამოყალიბებული და ფართოდ გავრცელებული, რაც განპირობებულია თავად დარგის სირთულითა და თავისებურებებით. სწორედ ამ ფაქტმა გადაგვაწყვეტინა პარალელური კორპუსლინგვისტიკის მეთოდოლოგიითა და კანონის ინტერპრეტაციის ტელეოლოგიური მიდგომების გამოყენებით კვლევისათვის დაგვემუშავებინა საქართველო-ევროკავშირის ასოცირების ხელშეკრულების 1-4 კარის 1-278 გვერდები და შეგვექმნა ქართულ-ინგლისური იურიდიული პარალელური კორპუსი. ეს მცდელობა კიდევ ერთი წინადადგმული ნაბიჯია თანამედროვე ქართულ კორპუსლინგვისტიკაში, რომელიც თავადაც ჩამოყალიბების ეტაპზეა. მას შეიძლება ვუწოდოთ მცირე მოცულობის იურიდიული პარალელური კორპუსი.

კორპუსის შექმნისას მრავალი გამოწვევის წინაშე ვდგებით, როგორებიცაა: თანამედროვე ტექნოლოგიების ფლობა, ადამიანური რესურსები, მაგრამ ყველაზე ძნელი მათ შორის მონაცემთა შეგროვებაა. საკმაოდ რთული იქნებოდა ორენოვანი ურთიერთშესატყვისი იურიდიული ტექსტების მოპოვება, თუ გავითვალისწინებთ მსგავსი სტილის კონფიდენციალურობასაც. საბოლოოდ არჩევანი შევაჩერეთ ასოცირების ხელშეკრულებაზე, რომელიც მრავალ ენაზეა თარგმნილი პროფესიონალი მთარგმნელების მიერ და ტექსტებიც არაკონფიდენციალურია,

პირიქით, ხელმისაწვდომია როგორც პროფესიონალების, ასევე ნებისმიერი დაინტერესებული პირისათვის.

სადისერტაციო ნაშრომის მოკლე აღწერილობა:

ა) **კვლევის მიზანი:** კვლევის მიზანია იურიდიული ტექსტების თარგმანის ანალიზის მოდელირების პრეცედენტის შექმნა პარალელურ კორპუსში.

შესაბამისად, რამდენიმე კვლევითი ამოცანა განვსაზღვრეთ:

1. თარგმანმცოდნეობის კვლევების მეთოდების ინტეგრირება კორპუსლინგვისტიკის თეორიისა და პრაქტიკის მეთოდებთან;

2. საკვლევი ორენოვანი ტექსტების პირველადი, შეპირისპირებითი არამანქანური ანალიზი, რომელიც რამდენიმე ურთიერთდაკავშირებულ პროცესს მოიცავს: ა. სინტაქსური, მორფოლოგიური, ტერმინოლოგიური თვისებრივი მონაცემების დადგენა, ბ. მთარგმნელობითი პრობლემების დადგენა; გ. ზემოთაღნიშნულ ანალიზზე დაყრდნობით სტრუქტურირებისა და მარკირების მოდელების განსაზღვრა;

3. საკვლევი ორენოვანი ტექსტების პარალელური კორპუსის ანალოგიური განთავსება და მანუალური მარკირება;

4. ტექნიკური ხასიათის სამუშაოს შესრულება თარგმანის ანალიზატორში მარკირებული ტექსტების პროცესირებისათვის;

5. ანალიზატორით მიღებული სტატისტიკური მაჩვენებლების, პარალელური სალექსიკონო მონაცემებისა და პარალელურად სტრუქტურირებული ტექსტების შემდგომი ანალიზი თარგმანმცოდნეობის მიმართულებით.

კვლევის მიზნების შესაბამისად კორპუსლინგვისტიკის თეორიული საფუძვლების დამუშავების მომდევნო ეტაპზე

შეიქმნა ემპირიული მასალის კორპუსი, რომელიც შედგენილია ქართულსა და ინგლისურ ენებზე არსებული იურიდიული ტექსტების პარალელური ტექსტებისაგან.

პარალელურ კორპუსში შეიძლება ერთდროულად განთავსდეს ტექსტი და თარგმნითი ვარიანტი, რამაც საშუალება მოგვცა შეგვექმნა იურიდიული ტექსტების თარგმანის ანალიზის ინდივიდუალური მოდელი ინგლისური და ქართული ტექსტების პარალელური კორპუსისათვის. კვლევა განვახორციელეთ ემპირიული მასალის წინასწარი დამუშავების შემდეგ ავტომატურ და ნახევრადავტომატურ რეჟიმში. კვლევისას გამოვიყენეთ პრაგმატიკული, სემანტიკური და ტერმინოლოგიური მიდგომები.

### **ბ) საკითხის აქტუალურობა, სიახლე და მნიშვნელობა:**

საკითხის აქტუალურობა, სიახლე და მნიშვნელობა განპირობებულია (1) კვლევის ობიექტითა და (2) კვლევის მეთოდოლოგიით.

მარკირებული და ანოტირებული ტექსტების ენობრივი რესურსების შექმნა ენის სწავლისა და კვლევისათვის ბოლო ათწლეულის განმავლობაში მთავარი სამეცნიერო მიმართულება გახდა. ენობრივი მასალების სამეცნიერო ანალიზისათვის ტექნოლოგიურად დამუშავებული კოლექციები ინფორმაციას გვაწვდის ბუნებრივი ენების შესახებ. ამ მხრივ პარალელური კორპუსი განსაკუთრებულ ინტერესს იწვევს. თარგმანის კორპუსი მკვლევარს კონკრეტული კვლევის საკითხის გადაწყვეტის საშუალებას უნდა აძლევდეს. სხვა სიტყვებით რომ ვთქვათ, არასპეციფიკური კორპუსის მკვლევარს შეუძლია შეარჩიოს და გამოიყენოს კორპუსის ტექსტები სხვადასხვა ტიპის შეპირისპირებითი კვლევისათვის.

საკანონმდებლო ტერმინების ადეკვატური თარგმანი წარმოადგენს საზოგადო პრობლემას, ვინაიდან თარგმანში ხშირია ტერმინთა კონცეპტუალური განსხვავება. ეს უკანასკნელი გულისხმობს იმას, რომ ტერმინების კონცეპტუალური განსხვავებები პოტენციურად შეიძლება გახდეს კანონის არასწორად ინტერპრეტაციის განმაპირობებელი მიზეზი.

ნაშრომის კიდევ ერთი სიახლე არის მონაცემების ტექნოლოგიური ბაზა, რომელიც შეიქმნა კვლევის პროცესში, როგორც ადაპტირებული ღია პლატფორმა სპეციალიზებული ენობრივი კორპუსებისათვის. ის უზრუნველყოფს ქართული ენის პარალელური კორპუსის განვითარებას ზოგადად. ასევე ის არის დაინტერესებული სპეციალისტებისათვის ერთიანი სამუშაო პლატფორმა, რომელიც აღჭურვილია კვლევის ჩვენეული მოდელით.

### **გ) კვლევის მეთოდები და მისი სიახლე**

არამანქანური ანალიზისათვის გამოვიყენეთ საკვლევი ტექსტების მხოლოდ ნაწილი (კარი 1-4, გვ.1-278). კვლევის მე-2 ამოცანის გ) პუნქტის შესაბამისად, დავადგინეთ მთარგმნელობითი პრობლემები საანალიზოდ შერჩეულ მონაკვეთზე და განვსაზღვრეთ მთარგმნელობითი უზუსტობების შემდეგი ოთხი კატეგორია:

1. კალკირება და მის შედეგად აზრობრივი ბუნდოვნების წარმოშობა, კერძოდ, სინტაქსური კონსტრუქციების კალკირება და ამის შედეგად გამოწვეული აზრობრივი ბუნდოვნება;
2. ცალკეული ლექსემის კალკირება;
3. ნასესხობების უმართებულო გამოყენება ქართული შესატყვისების ნაცვლად;

4. საკანონმდებლო ტერმინების თარგმანის ადეკვატურობა.

შესაბამისად, მათი ანალიზი წინამდებარე კვლევაში ჩვენი ქვე-ამოცანა გახლავთ.

კორპუსლინგვისტიკის თეორიული საფუძვლების დამუშავების მომდევნო ეტაპზე თემის ხელმძღვანელის მიერ ქართული და ინგლისური ტექსტების მანქანური ანალიზისთვის შედგენილი ლინგვისტიკური მარკირების TagSet -დან ჩვენი კვლევის ამოცანების შესაბამისი მარკირების ერთეულები განვსაზღვრეთ.

კვლევისათვის განისაზღვრა მეთოდოლოგია შემდეგი კლასიფიკაციით:

- 1) საკვლევი ლინგვისტიკური და მთარგმნელობითი ერთეულების დადგენა - მანუალური დამუშავების პროცესის დროს;
- 2) სტატისტიკური ანალიზის მოდელის განსაზღვრა - მანუალური დამუშავების პროცესის დროს;
- 3) ტერმინების პროგნოზული ანალიტიკა - კორპუსის პროგრამული ანალიზატორით მიღებული შედეგების საფუძველზე ;
- 4) თარგმანის ხარისხობრივი ანალიზი - კორპუსის პროგრამული ანალიზატორით მიღებული შედეგების საფუძველზე.

მომდევნო ეტაპზე კვლევისათვის შერჩეული ორენოვანი ტექსტების მანუალური დამუშავება დავიწყეთ, რაც შეიძლება აღვწეროთ, როგორც მანქანური ანალიზისათვის მოსამზადებელი წინასწარი ანალიტიკური პროცესი. მარკირებიული ტექსტები განთავსდა პროგრამულ ანალიზატორში, მიღებული მონაცე-



მების საფუძველზე სამეცნიერო კვლევა თარგმანმცოდნეობის მიმართულებით ჩავატარეთ.

**ენა და თარგმანი:** გლობალიზაციის ერაში ენა ძირითადად თარგმანის საშუალებით ვითარდება. ბუნებრივია, ბოლო ოცი წლის განმავლობაში იგი განიცდის ისეთი ახალი კონცეპტებისა და მათი აღმნიშვნელი ტერმინების ზეგავლენას, როგორც მანამდე არ არსებობდა არამართო ქართული ენის თეზაურუსში, არამედ ჩვენი მსოფლალქმის ცნებათა ველში. თუ გავითვალისწინებთ, რომ ახალ კონცეპტთა მართებული დამკვიდრება მთარგმნელთა მისიაა, მათ უდიდესი პასუხისმგებლობა ეკისრებათ ენის განვითარებაში, რადგან მათი აღმნიშვნელი ტერმინები მხოლოდ მათ გადააქვთ ჩვენს ცნებათა ველში და ისინი ამკვიდრებენ ჩვენს თეზაურუსში. თუ ამ პროცესს წარმოვიდგენთ, როგორც ენისა და ენაში დამკვიდრებული ახალი კონცეპტების გამოყენების უწყვეტ სოციალურ ჯაჭვს, მათ მიერ თარგმნილი მასალა ჯაჭვური პრინციპით ხვდება მასმედიისა და პოლიტიკოსების დისკურსში, ხოლო ეს უკანასკნელნი სოციალური მედიასაშუალებების გამოყენებით ამკვიდრებენ სოციუმის დანარჩენი წევრებისათვის არა მხოლოდ გაუმართავ და ხშირად გაუგებარ ტერმინებს, არამედ კალკირებულ, გაუმართავ სინტაქსურ კონსტრუქციებსაც.

**ნაშრომის სტრუქტურა და მოცულობა:** ნაშრომი მოიცავს შესავალს, ოთხ თავსა და თერთმეტ ქვეთავს, დასკვნას, გამოყენებულ ლიტერატურას, ინტერნეტ წყაროებსა და დანართებს. ნაშრომი 133 ნაბეჭდი გვერდია.

## ნაშრომის მოკლე შინაარსი

თავი I. კორპუსლინგვისტიკა, როგორც ინოვაციური სამეცნიერო მიმართულება -ორი ქვეთავისგან შედგება:

1.1. კორპუსლინგვისტიკის გაგებისთვის - ქვეთავში განმარტებულია კორპუსლინგვისტიკის, როგორც ჰუმანიტარული მეცნიერების შედარებით ახალი მიმართულების რაობა, რომელიც მეცნიერების ორი დარგის: ჰუმანიტარიისა და ინფორმატიკის შერწყმის საფუძველზე წარმოიშვა. იგი, როგორც ერთგვარი „სიმბიოზი“, საკმაოდ სწრაფად განვითარდა თეორიული ლინგვისტიკისა და ინფორმატიკის განვითარების პარალელურად და სულ მალე ინტერდისციპლინურ სამეცნიერო დარგად ჩამოყალიბდა. სამეცნიერო წრეებში აზრთა სხვადასხვაობას იწვევს კითხვა: არის თუ არა კორპუსლინგვისტიკა დამოუკიდებელი სამეცნიერო დისციპლინა და თუ ეს ასეა, რა განსხვავებაა კორპუსლინგვისტიკასა და ტრადიციულ ლინგვისტიკას შორის.

კამათის საფუძველს იძლევა ის ფაქტი, რომ კორპუსლინგვისტიკა საერთაშორისო კლასიფიკაციაში (DDC – Dewey Decimal Classification, სამეცნიერო დარგების სისტემატიზაციის ერთერთი ყველაზე გავრცელებული საერთაშორისო კლასიფიკაცია, შემუშავებული ამერიკელი მკვლევარის მელვილ დივეის მიერ და UDC – Universal Decimal Classification – სამეცნიერო დარგების სისტემატიზაციის ევროპული კლასიფიკაცია (შექმნილი ბელგიელი მკვლევარების პაულ ოთლეტის და ჰენრი ფონტენის მიერ 1895 წელს) არ არის ასახული, როგორც დამოუკიდებელი სამეცნიერო დისციპლინა. „ლინგვისტიკას საკუთარი კოდი აქვს (DDC-ში 4[00], UDC-ში კი 8), ხოლო ინფორმატიკას – საკუთარი (004 DDC- შიგ და UDC-შიც )“ (თანდაშვილი, 2014:13).

კორპუსლინგვისტიკა დღესდღეობით შედარებით ახალი მიმართულებაა და ჯერ კიდევ კვლევის და ჩამოყალიბების პროცესშია, ამიტომ მისი განსაზღვრებებიც არაერთგვაროვანია. ჩვენ განვიხილეთ სხვადასხვა მკვლევრის მიერ შემოთავაზებული განმარტებები, რომლებიც განსხვავებული პოზიციებიდან გამომდინარეა ჩამოყალიბებული. შევეცადეთ, ჩაგვეტარებინა რამდენიმე მათგანის შეპირისპირებითი ანალიზი, რათა ჩვენთვის შესაძლებელი გამხდარიყო ჩვენი შეხედულებისამებრ, ყველაზე მისაღები და შედარებით ამომწურავი განმარტების შეჯერება. საგულისხმოა ის ფაქტიც, რომ კორპუსლინგვისტიკაში თეორიული შრომების უმრავლესობის შესწავლისას მივედით იმ დასკვნამდე, რომ ისინი პრაქტიკოსი კორპუსლინგვისტების მიერაა დაწერილი, რაც დასტურია იმისა, რომ ინტერდისციპლინურ მეცნიერებებში წაშლილია ზღვარი თეორიასა და პრაქტიკას შორის. თუკი ადრეულ პერიოდში ლინგვისტური კვლევები ძირითადად თეორიით შემოიფარგლებოდა და მხოლოდ შემდეგ ინერგებოდა ისინი პრაქტიკაში (ანუ პროცესი „ზემოდან – ქვემოთ“), კორპუსლინგვისტიკის თეორიული შრომები უშუალოდ პრაქტიკულ კვლევებს ეყრდნობა (პროცესი „ქვემოდან – ზევით“). უფრო მეტიც, თეორიულად დამტკიცებული მასალა კვლავ პრაქტიკული გამოყენებისთვისაა განკუთვნილი („ქვემოდან – ზევით – ქვევით“).

დავუბრუნდეთ ცალკეული მკვლევრის მიერ შემოთავაზებულ კორპუსლინგვისტიკის განმარტებებს და მათ ანალიზს:

მკვლევართა ნაწილი კორპუსლინგვისტიკას განიხილავს, როგორც მეთოდს, რომელიც საშუალებას გვაძლევს ლინგვისტური პრობლემები დიგიტალურ ტექსტებზე დაყრდნობით ვიკვლიოთ. კვლევის განსახორციელებლად საჭიროდ მიიჩნევა ენის წერილობითი და ზეპირმეტყველების ნიმუშების გაციფ-

რება და მათი ანოტირება მეტყველების ნაწილების მიხედვით. უმეტესწილად კვლევის ზოგი ასპექტი მოითხოვს დიდი მოცულობის მონაცემების გადამუშავებას, რომელიც უფრო დეტალურ, ავტომატურ ანოტაციას საჭიროებს (კომპიუტერული ლინგვისტიკის პორტალი <http://www.computerlinguistik.org/portal/portal.html?s=Korpuslinguistik>).

კორპუსლინგვისტიკის ზემოთმოყვანილი განმარტებებიდან ნათლად ჩანს, თუ რამდენად ძნელია, შევჯერდეთ მის ცალსახა დეტერმინაციაზე, თუმცა მათზე დაყრდნობით შეგვიძლია თავი მოვუყაროთ ცალკეულ პუნქტებს, რომლებიც, თავის მხრივ, განაპირობებენ კორპუსლინგვისტიკის მიზანს, კერძოდ:

კორპუსლინგვისტიკა, უპირველეს ყოვლისა, ეყრდნობა ავტენტურ ენობრივ რესურსებს, 2) გაციფრებულ რესურსებს, 3) იყენებს დიდი მოცულობის ემპირიულ მასალას.

**1.2.კორპუსის რაობა** - კორპუსლინგვისტიკაზე საუბრისას მიზანშეწონილად მიგვაჩნია გავერკვეთ თავად კორპუსის რაობაში. სიტყვა „corpus“ (მრ. „corpora“) ლათინური წარმოშობისაა და ქართულად „სხეულს“ ნიშნავს. თითქოს უცნაურიცაა, რა საერთო შეიძლება ჰქონდეს მას ლინგვისტიკასთან. თუ ჩავუღრმავდებით, „კორპუსი“ ამ შემთხვევაში ასოცირდება ერთ მთლიანობასთან. შემთხვევითი არ არის ის ფაქტიც, რომ ისევე როგორც სხეული ასოცირდება ცოცხალ ორგანიზმთან, კორპუსიც წარმოადგენს ენის გაციფრებულ ელემენტთა კინეტიკურ ერთობლიობას და არა უბრალოდ ენობრივ ელემენტთა უმოქმედო, „უსიცოცხლო“, სტატიკურ ერთობლიობას.

განვიხილავთ ტერმინს „კორპუსი.“ მას რამდენიმე მნიშვნელობა აქვს ინგლისურ ენაში. ენათმეცნიერებაში გვესმის ტერმინი „კორპუსი“, როგორც „ტექსტების კრებული, რომელიც

გამოყენებული იქნება ლინგვისტური ანალიზისთვის. „კორპუსში დაფიქსირებული ენა ბუნებრივია“ (ტოგენინი - ბონელი, 2001: 2). სინკლერმა კორპუსი განსაზღვრა, როგორც „ბუნებრივად წარმოქმნილი ენობრივი ტექსტის კრებული, რომელიც დამახასიათებელია სახელმწიფოსთვის ან ენის მრავალფეროვნებისთვის“ (Sinclair, 1991: 171). 2005 წელს მან ჩამოაყალიბა კორპუსის უფრო ზუსტი განსაზღვრება, რადგან „კორპუსი არის ტექსტების ერთობლიობა ელექტრონული ფორმით, რომელიც შეირჩა გარე კრიტერიუმების შესაბამისად და შეძლებისგვარად წარმოადგენს ენის მრავალფეროვნებას, როგორც მონაცემთა წყაროს“ (Sinclair, 1991: 171).

მსგავსად ნებისმიერი კვლევისა, კორპუსული კვლევის დროს უპირველესი და გადამწყვეტი მნიშვნელობა ენიჭება მეთოდებს. გამომდინარე იქიდან, რომ კორპუსი ეყრდნობა დიდ მონაცემებს (Big Data), მათი დამუშავება შეუძლებელია კვლევის ტრადიციული მეთოდების გამოყენებით, ეს აღემატება მრავალი მკვლევრის ერთობლივ ინტელექტუალურ შესაძლებლობებსაც კი. აუცილებელი ხდება სპეციალური კომპიუტერული პროგრამების გამოყენება. კორპუსლინგვისტები (ს.შრაიბმანი, დ.ბერი, მ.თანდაშვილი, რ.ხალვაში და სხვანი) დიდ მონაცემებს სხვადასხვა ნიშან-თვისების მიხედვით ახასიათებენ, რომელთაგან კორპუსული კვლევისათვის განსაკუთრებით მნიშვნელოვნად მიგვაჩნია შემდეგი: მოცულობა (volume), სიჩქარე (velocity) - მონაცემთა დამუშავების სიჩქარე), მრავალფეროვნება (variety), უტყუარობა (veracity) და ცვალებადობა (variability).

იგივე კორპუსლინგვისტები აღიარებენ დიდი მონაცემების ანალიზის სხვადასხვა მეთოდს, რომელთაგან განსაკუთრებით გამოვყოფდით შემდეგს: მონაცემთა მოპოვება (Data Mining), თანაშემკრებლობა (Crowdsourcing) – მონაცემთა კატეგორ-

რიზაცია და გამდიდრება მრავალი ადამიანის მონაწილეობით საჯარო შეთავაზების საფუძველზე; სხვადასხვა წყაროს მონაცემთა შერწყმა და ინტეგრაცია (Data Fusion and Integration), მანქანური შესწავლა (Machine Learning), მოდელების ამოცნობა (Pattern Recognition), სიმულაციური მოდელირება (Simulation Modeling), სივრცობრივი ანალიზი (Spatial Analysis), სტატისტიკური ანალიზი, ანალიტიკური მონაცემების ვიზუალიზაცია.

წინამდებარე ნაშრომში, უპირველეს ყოვლისა, ვიყენებთ მონაცემთა მოპოვების (**Data Mining**) მეთოდს. საგულისხმოა, რომ ტერმინი **data mining** (რომელიც ეკუთვნის გ.პიატეტსკი-შაპიროს, 1989 წ.), დიდი მონაცემების მსგავსად (**Big Data**), წიაღისეულის ანალოგიით შეიქმნა. ინგლისურიდან თარგმნისას **mining** მოპოვებას ნიშნავს. მართლაც, კოლოსალური მოცულობის ინფორმაციიდან მოიპოვებენ კვლევითვის საჭირო მასალას. მონაცემთა მოპოვების გზით შესაძლებელია ადრე ნაკლებად შესწავლილი, ხშირ შემთხვევაში უცნობი, პრაქტიკულად გამოყენებადი და ინტერპრეტაციისათვის ხელსაყრელი მონაცემების მოძიება, გადამუშავება და გამოთვლა. ჩვენი მიზანია, არა უკვე არსებული ჯგუფების კლასიფიკაცია, არამედ ახალი ჯგუფებისა და კანონზომიერებების გამოვლენა დიდ მონაცემთა ბაზებში. ეს გახლავთ მონაცემთა ბაზებში ცოდნის აღმოჩენის პროცესის მანქანური დამუშავების ეტაპი. ეს უკანასკნელი კი ჰიპოთეზების განერიერების, ვერიფიცირების და დაზუსტების საშუალებას გვაძლევს.

მონაცემთა მოპოვებისას ვეყრდნობით თანამედროვე ციფრულ მეთოდებსა და ტრადიციულ თეორიულ დისციპლინებს, რომლებიც ფართოდ გამოიყენება პრაქტიკოსი თუ თეორეტიკოსი კორპუსლინგვისტების მიერ, როგორებიცაა: მონაცემთა ბაზები (**Database**) – ელექტრონული ინფორმაციის, ციფრული

ობიექტების ერთობლიობა, რომელთა მეშვეობითაც მონაცემებს ავტომატურ რეჟიმში ვეძიებთ და დავამუშავებთ; ხელოვნური ინტელექტი (**Artificial Intelligence**) – ავტომატური სისტემის თვისება, რომელსაც გააჩნია ადამიანური ინტელექტის გარკვეული ფუნქციები; მანქანური შესწავლა (**Machine Learning**) – სპეციალურად შემუშავებული კომპიუტერული ალგორითმები, რომლებსაც აქვთ მუშაობის პროცესში ავტომატურად გაუმჯობესების (**დაჭკვიანების**) უნარი, რაც იმთავითვე პროგრამირებულია მასში; **ალგორითმიზაცია** – წინასწარ შედგენილი ალგორითმების გამოყენება დასახული ამოცანის კომპიუტერის მეშვეობით გადაჭრის მიზნით.

მონაცემთა მოპოვება საკმაოდ ხანგრძლივი და შრომატევადი პროცესია, რომელიც რამდენიმე ეტაპს მოიცავს. კორპუსლინგვისტები განასხვავებენ სხვადასხვა ეტაპს, რომელთაგან გამოვყოფდით შემდეგ ძირითადს: თავისუფალი ძიება (**Discovery**), რომლის დროსაც გამოვლინდება გარკვეული კანონზომიერებები დედუქციური მეთოდით – კერძოდან ზოგადისაკენ; **ვალიდაცია** – იმის დადასტურება, რომ შევქმენით სწორი პროდუქტი; **ვერიფიცირება** – ამ ეტაპზე ვადასტურებთ, რომ შევქმენით ისეთი პროდუქტი, როგორსაც ვგეგმავდით; პროგნოსტიკური მოდელირება (**Predictive Modeling**) – აღნიშნულ ეტაპზე გამოვლენილი კანონზომიერებები გამოიყენება უცნობი მნიშვნელობის პროგნოზირებისათვის; გამონაკლისების ანალიზი (**Forensic Analysis**) – აღნიშნულ ეტაპზე ვახდენთ კანონზომიერებებში მიკვლევული ანომალიების გამოვლენას და მათ ახსნას (ხალვაში, 2018:38). წინამდებარე ნაშრომი სწორედ მონაცემთა მოპოვების აღნიშნულ ეტაპებს მოიცავს, რომელთა გამოწველივით ანალიზს და საბოლოო შედეგს მომდევნო თავებში შემოგთავაზებთ.

მონაცემთა მოპოვების შემდეგი, არანაკლებ შრომატევადი და ხანგრძლივი პროცესია ტექსტის კოდირება, ანუ მოპოვებულ მონაცემთა სტრუქტურირება. კორპუსში მონაცემთა სტრუქტურირების ყველაზე გავრცელებულ საშუალებად მარკირება მიიჩნევა. **მარკირება** (mark-up) შეიძლება განვმარტოთ, როგორც **მონიშვნა**, შესაბამისად, იგი გულისხმობს თეგების გამოყენებას, რომლებიც დაერთვის სიტყვებს და ფრაზებს. თეგების (tag) სხვადასხვა ფორმა გამოიყენება. ვიზუალურად თეგები შეიცავს ასობგერებსა და გარშემორტყმულ სიმბოლოებს კუთხური ფრჩხილების სახით (<...>). თეგები, როგორც წესი, წყვილებადაა წარმოდგენილი, მაგალითად: <n> და </n>. პირველი განიხილება, როგორც საწყისი თეგი, ხოლო მეორე – როგორც სასრული. საწყის და სასრულ თეგებს შორის თავსდება ტექსტი – საანალიზო მასალა, რომელსაც შიგთავსის სახელწოდებით მოიხსენიებენ. შიგთავსი თეგებში გვხვდება როგორც დიდი, ასევე პატარა ასობგერებით ჩაწერილი, მაგალითად: <n > ან <N >, თუმცაღა ეს არსებითად არ განასხვავებს მათ ერთმანეთისაგან. თეგები, როგორც წესი, იერარქიული სტრუქტურის მქონეა და ისინი ტექსტშია ჩაშენებული. ჩვენი კვლევისას შევიმუშავეთ სპეციალური თეგები, რომელთა გამოყენებასაც დაეფუძნა საანალიზო ტექსტების მარკირება წინასწარ შექმნილი პროგრამული ანალიზატორისთვის.

შესაბამისად, ნებისმიერი მარკირება შეგვიძლია განვიხილოთ, როგორც ტექსტის ინტერპრეტირების საშუალება. ტექსტში გარკვეული ჩარევის გზით მასში დამატებით კოდირებულ ინფორმაციას ჩავტვირთავთ, რაც, თავის მხრივ, საშუალებას გვაძლევს, გავანალიზოთ იგი, განვახორციელოთ კვლევა, შევადაროთ იგი სხვა ტექსტებს. აქედან გამომდინარე, მარკირება კორპუსლინგვისტიკის მნიშვნელოვანი ელემენტია, ვინაიდან



იგი წარმოადგენს ტექსტის სტრუქტურირების ძირითად საშუალებას.

როგორც აღვნიშნეთ, თეგირება დროში საკმაოდ გაწელილი, შრომატევადი და, იმავდროულად, საინტერესო პროცესია. იგი რამდენიმე ეტაპად მიმდინარეობს.

ზემოთქმული მოსაზრებების გათვალისწინებით თეგირების პირველ ეტაპად შეგვიძლია მივიჩნიოთ სქემისა და მოდელის შერჩევა დასმული ამოცანის შესაბამისად, ანუ იმის გათვალისწინებით, თუ რა სახის ანალიზი გვსურს ტექსტში, როგორებიცაა, მაგალითად, კულტურული თავისებურებები, გენდერული ბალანსი თუ, ჩვენ მიერ შერჩეული საკვლევი თემის შემთხვევაში, თარგმნის ეკვივალენტობის პრობლემა პარალელურ ტექსტებში. თეგების ნაკრები უნდა შეირჩეს იმის გათვალისწინებით, თუ რამდენად შეესაბამება ის როგორც ტექსტის თემას, ასევე დასახულ კვლევით სამუშაოს. მაგალითად, ჩვენ მიერ ჩასატარებელი კვლევის შემთხვევაში, კორპუსში შესატანი მასალის ავტომატურ და ნახევრადავტომატურ რეჟიმში დამუშავების დროს თეგირებას ვახორციელებთ სამ ეტაპად: 1. მორფოლოგიური ანალიზი, 2. სინტაქსური ანალიზი, 3, სემანტიკური ანალიზი. მომდევნო თავებში განხილულია კონკრეტულ მაგალითები, მათი ანალიზი და დასკვნები.

**II თავი. იურიდიული პარალელური ტექსტების კვლევის აპრობირებული მეთოდოლოგია უცხოეთის მოდელეზზე დაყრდნობით - შედგება ოთხი თავისაგან.**

**2.1. ევროკომისიის თარგმანის გენერალური დირექტორატის თარგმანის ხარისხის სტანდარტები და მეთოდური რეკომენდაციები** - ვინაიდან ჩვენი სადისერტაციო ნაშრომის კვლევითი მიმართულება საქართველოს ევროკავშირთან ასოცირების

ხელშეკრულების ინგლისური და ქართული ტექსტები და მათი მთარგმნელობითი ეკვივალენტობის ანალიზი გახლავთ, მიზანშეწონილად მიგვაჩნია, დაწვრილებით გავეცნოთ **თარგმანის ხარისხის სტანდარტებსა და მეთოდურ რეკომენდაციებს**, რომლებიც ევროკომისიის თარგმანის გენერალური დირექტორატის მიერ არის ჩამოყალიბებული და რაც მომავალში, ასოცირების შემდეგ, გახდება სავალდებულო სახელმძღვანელო-გზამკვლევი საქართველოში თარგმანის ხარისხის სტანდარტიზაციისათვის;

ევროკომისიის თარგმანის გენერალური დირექტორატი აწესებს ხარისხის მართვის ჩარჩოებს, რომლის მიხედვითაც თარგმანი უნდა აკმაყოფილებდეს უშუალო დამკვეთის, ევროკომისიის სხვა პარტნიორი ორგანიზაციებისა თუ საბოლოო მომხმარებელთა მოთხოვნებს და ამართლებდეს მათ მოლოდინს (DGT Translation Quality Guidelines, 2015:1). შემუშავებულია სპეციალური მეთოდური რეკომენდაციები, რომლებიც განსაზღვრავენ სათარგმნი ტექსტის მიზნებს, მოსალოდნელ რისკებსა და ხარისხის სპეციფიკურ მოთხოვნებს. თვდ თავადვე სთავაზობს დახმარებას მთარგმნელებს და რედაქტორებს, რათა უზრუნველყონ მიზანთან შესატყვისობა.

ყველა ოფიციალურ ენას აქვს ერთნაირი ოფიციალური სტატუსი, მოთხოვნები და ამოცანები, რომლებიც თითოეულ მათგანზე თანაბრად ვრცელდება. ეს ნიშნავს, რომ 24 ოფიციალური ენიდან 23 შემთხვევაში მოთხოვნები და ამოცანები უნდა შესრულდეს თარგმანის მეშვეობით, ამგვარად, ტექსტის „დეფაქტო“ შემდგენელი გახლავთ მთარგმნელი.

მსგავსი დონის თარგმანი მოიხსენიება, როგორც „ინსტიტუციონალური თარგმანი“ ან „მულტიენობრივი კანონწარმოება“ (DGT Translation Quality Guidelines, 2015:2). შესაბამისად, მთარგმნელები თავიანთი ენების „ინსტიტუციონალურ ხმას“

წარმოადგენენ და თანაბრად ავთენტურ ტექსტებს ქმნიან. ამგვარად, არსებული მეთოდური რეკომენდაციები განკუთვნილია არა მხოლოდ წყარო-ტექსტების შემდგენელთათვის, არამედ მთარგმნელებისა და თარგმანის გამსწორებლებისთვისაც. ხარისხის უმთავრესი მოთხოვნაც სწორედ ისეთი ტექსტების შექმნა გახლავთ, რომლებიც ყველა ენაზე წყარო ტექსტის მსგავსად იკითხება.

თუმცაღა, თუ ჩავუღრმავდებით, თარგმანის აბსოლუტური იგივეობის მიღწევა შესაძლებელია მხოლოდ ზღვრულად, ვინაიდან 24-დან 23 ტექსტი არის თარგმანი და ეყრდნობა წყარო ტექსტს. მაშასადამე, საჭიროა განვსაზღვროთ ამ დამოკიდებულების ხარისხი, თუ რას ნიშნავს თარგმანის გენერალური დირექტორატის რეკომენდებული დედნისა და თარგმანის იგივეობის ფორმულა: „იმდენად, რამდენადაც შესაძლებელია“ პრაქტიკაში გამოყენებისას. პასუხს ფაქტობრივად განსაზღვრავს ტექსტის მიზანი. გარდა ამისა, განსხვავებულ ენაში ის შეიძლება სხვადასხვა იყოს, ვინაიდან მეთოდური მითითებები ერთი და იგივე არ არის განსხვავებულ ტექსტებსა და ენებში.

თარგმანის გენერალური დირექტორატი განასხვავებს ორი სახის ხარისხის კონტროლს (DGT Translation Quality Guidelines, 2015:35):

- **შემოწმება** (revision), რომელიც გულისხმობს სამიზნე და წყარო ტექსტების ორენოვან კვლევას დასახულ ამოცანებთან მათი შესატყვისობის განსაზღვრის მიზნით;
- **რევენზირება** (review), რაც გულისხმობს სამიზნე ტექსტის ერთენოვან კვლევას დასახულ ამოცანებთან მისი შესატყვისობის განსაზღვრის მიზნით.

ჩაუღრმავდეთ მეთოდურ მითითებებს სამართლებრივი დოკუმენტაციისათვის, რომლებიც შეეხება შემდეგი სახის დოკუმენტებს:

1. ევროკავშირის სამართლებრივი აქტები: ხელშეკრულებები, რეგულაციები, დირექტივები, გადაწყვეტილებები, რეკომენდაციები, მოსაზრებები, საერთაშორისო შეთანხმებები;

2. დოკუმენტები, რომლებიც გამოიყენება ადმინისტრაციულ და სასამართლო საქმისწარმოებასა და მოკვლევაში, როგორებიცაა: სამართალდარღვევა, ანტიმონო-პოლიური, სახელმწიფო სუბსიდირებისა თუ სხვა საქმეები, მაგალითად: სასამართლო სარჩელი, ოფიციალური წერილები, შემაგებებელი სარჩელი და სხვა;

3. შესყიდვების დოკუმენტაცია და დაფინანსების პროგრამები, ტენდერები, საგრანტო განაცხადებები და კონტრაქტები;

4. დასაქმების ცნობები, საკონკურსო და ტესტირებების დოკუმენტაცია.

**ყოველი თარგმნილი ტექსტი უნდა შეესაბამებოდეს საერთაშორისო სტანდარტებით (ISO 17100) განსაზღვრულ ძირითად პრინციპებსა და პროფესიონალური თარგმანის ხარისხის მოთხოვნებს.**

რაც შეეხება ევროკავშირის სამართლებრივ აქტებს, მათ აქვთ სამართლებრივი ძალა, კერძოდ: განსაზღვრავენ უფლებებს, ვალდებულებებს, კანონის მოსალოდნელ შედეგებს. მკითხველი, იქნება რიგითი მოქალაქე თუ სასამართლოს წარმომადგენელი, სრულიად დარწმუნებული უნდა იყოს დოკუმენტში არსებული ინფორმაციის სანდოობაში და იმოქმედოს შესაბამისად. **სამართლებრივი აქტების ყველა ენაზე არსებული ვერსია თანაბრად ავთენტური უნდა იყოს.** ისინი ყველა ენაზე უნდა შეიცავდნენ ერთსა და იმავე შინაარსს და ჰქონდეთ ერთ-

ნაირი სამართლებრივი მოქმედების ძალა. დაუშვებელია არსებობდეს მეტ-ნაკლები უპირატესობა სხვადასხვა მოქალაქის, კომპანიისა თუ წევრი სახელმწიფოებისათვის, რაც განპირობებული იქნება თარგმანში დაშვებული შეცდომებითა თუ სხვადასხვა ენაზე არსებულ ვარიანტებს შორის შეუსაბამობით.

ტერმინოლოგია უნდა იყოს ორმხრივად შესატყვისი, კერძოდ, ერთი და იგივე ტერმინი სინონიმების და დამატებითი ფორმულირებების გარეშე უნდა გამოიყენებოდეს როგორც სამართლებრივ აქტში, ასევე ნებისმიერ ძირითადასა თუ პარალელურ აქტში.

ახალი ტერმინის შექმნა დასაშვებია მხოლოდ კონცეპტუალური ანალიზის შედეგად, რასაც ტერმინისა თუ ფრაზის ტერმინოლოგიურ გაგებასთან მივყავართ.

რაც შეეხება ტერმინის კონცეპტუალურ არეალს, ის უცვლელი უნდა დარჩეს. ევროკავშირის კანონმდებლობის სპეციფიკური ბუნების გათვალისწინებით ცალკეული ერის სამართლებრივი სისტემისათვის დამახასიათებელი კონცეპტებისა და ტერმინოლოგიის გამოყენებისას სიფრთხილე უნდა გამოვიჩინოთ.

გასაგებად და სწორად თარგმნილი კანონები მოქმედი დემოკრატიის უტყუარი წინაპირობაა. ყველა ენაზე არსებული ვერსიები უნდა შეესაბამებოდეს ზოგადსაკანონმდებლო პრინციპებს, როგორებიცაა: კანონის წინაშე თანასწორუფლებიანობა (კანონი თანაბრად ხელმისაწვდომი და გასაგები უნდა იყოს ყველასთვის) და კანონის სანდოობა (შესაძლებელი უნდა იყოს მოსალოდნელი შედეგების წინასწარ განჭვრეტა) (DGT Translation Quality Guidelines; 2015: 6).

**2.2. Acquis Communautaire – პარალელური კორპუსი 20+ ენით** - ჩვენი კვლევის მიზნებისა და ამოცანების შესაბამისად, მიზანშეწონილად მივიჩნით ასოცირების ხელშეკრულების კორპუსი Acquis Communautaire განხილვა, რადგან საქართველოს ევროკავშირთან ასოცირების ხელშეკრულება, როგორც იურიდიული დოკუმენტი, სწორედ მისი ნაწილი გახდება; ხოლო დისერტაციის ფარგლებში ჩვენი კვლევის ობიექტია სწორედ აღნიშნული ასოცირების ხელშეკრულების ინგლისური და ქართული ტექსტები და მათი მთარგმნელობითი ეკვივალენტობის ანალიზი.

Acquis-ი აერთიანებს ევროკავშირის ხელშეკრულებებსა და კანონებს (დირექტივებს, დებულებებს, გადაწყვეტილებებს), დეკლარაციებსა და რეზოლუციებს, საერთაშორისო შეთანხმებებსა და სასამართლოს განაჩენებს. ის ასევე მოიცავს ევროკავშირის წევრი სახელმწიფოების მთავრობების ერთობლივ აქტებს სასამართლო და საშინაო საკითხებში (თავისუფლება, უსაფრთხოება და სასამართლო), ასევე საგარეო და უსაფრთხოების პოლიტიკის ერთობლივ აქტებს.

დღესდღეობით Acquis-ი თარგმნილია ევროკავშირის 22 ოფიციალურ ენაზე, რომლებიც წყარო ტექსტთან ერთად არსებობენ, როგორც პარალელური ტექსტები.

დღესდღეობით კი, რამდენადაც ჩვენთვის ცნობილია, არსებულ მრავალენოვან კორპუსებს შორის ყველაზე მასშტაბური **JRC-Acquis-ის**, რომელიც ოცზე მეტ ენას მოიცავს. ის ხელმისაწვდომია XML ფორმატში და შეიცავს მონიშნულ ტექსტებსა და თანაფარდობით ინფორმაციას 190-ზე მეტი ენობრივი წყვილისათვის.

JRC-Acquis-ის დამატებითი თვისება გახლავთ ის, რომ კლასიფიკაციის სპეციალური სისტემის - EUROVOC-ის თეზაუ-

რუსის გამოყენებით (კლასიფიკაციის სისტემა, რომელიც 6000-ზე მეტ იერარქიულად დალაგებულ ჯგუფს აერთიანებს), შესაძლებელი ხდება ტექსტების კლასიფიკაცია საგნობრივი მახასიათებლების მიხედვით. ტექსტის საგნობრივი მახასიათებლების ცოდნის შემთხვევაში შესაძლებელი ხდება სპეციფიკური ტერმინოლოგიის ჩამონათვალის შექმნა, ასევე დოკუმენტების კლასიფიკაციის პროგრამული უზრუნველყოფისა და ავტომატური ინდექსირების სისტემის გამოცდა.

**2.3. ასოცირების ხელშეკრულების თარგმანი სერბეთში -** წინამდებარე ნაშრომში მიზანშეწონილად მიგვაჩნია განვიხილოთ ის კორპუსლინგვისტიკური მეთოდები, რომლებიც ევროკავშირის სხვადასხვა ქვეყანაში მკვლევრებმა გამოიყენეს ასოცირების ხელშეკრულების თარგმანისა და კორპუსლინგვისტიკური დამუშავების მიზნით; კერძოდ, ჩვენი კვლევისათვის მიზანშეწონილია განვიხილოთ მ. ვუკჩევიჩის კვლევითი სტატია სათაურით: ევროკავშირის დირექტივები და სტანდარტები თარგმანისა და თარჯიმნობის შესახებ და მათი გამოყენება სერბეთში, ევროკავშირის არაწევრ ქვეყანაში, რომელიც სტაბილიზაციისა და ასოცირების პროცესს გადის. მ. ვუკჩევიჩი წერს<sup>1</sup>,

---

<sup>1</sup> “A wrong translation costs the citizens 20 million Euros. The ignorance of the employees in the National Health Insurance Fund” This news headline (Todorović 2013: 2) recently has aroused tempers. The director of the Serbian National Health Insurance Fund was taken to task for initiating the use of new health insurance cards despite the costs it will incur for the citizens at a time when the global financial crisis burst in upon Serbia. It is claimed that the health card, as it is called in translation, refers not to the health insurance card but to the health record. Although the translation from English to Serbian is obviously not wrong, confusions arose from the interference between the meanings of the European Health Insurance Card (or EHIC) and the electronic health record.

რომ ევროკავშირთან ასოცირების ხელშეკრულებაში **მთარგმნელობითი შეცდომა ქვეყანას 20 მილიონი ევრო დაუჯდა.**

მივიჩნით, რომ აღნიშნული ფაქტი ჩვენს თემაში დაწვრილებით ანალიზს საჭიროებდა, რათა (1) მხედველობაში მივიღოთ სერბეთის გამოცდილება და თავიდან ავირიდოთ სავალალო შემთხვევა, ფინანსური პრეცედენტი; (2) კიდევ ერთხელ დავსვათ საკითხი, იურიდიული დოკუმენტაციის ტერმინოლოგიის თარგმნის სწორი სტრატეგიის შემუშავების აუცილებლობაზე;

როგორც სერბეთის მაგალითზე აღმოჩნდა, თარგმანის რედაქტირების თანმიმდევრობა იწვევს პრობლემებს, რასაც მკვლევარ ბერტლოოტის ანალიზიც მოწმობს (Berteloot, 1999:101). ასოცირების ხელშეკრულებაზე მუშაობის პროცედურის არასახარბიელო შედეგები, კერძოდ, თარგმანის დაბალი ხარისხი მეცნიერთა მოსაზრებით გახლავთ ის, რომ **ლინგვისტური რედაქტირება** წინ არ უსწრებდა **სამართლებრივ რედაქტირებას**, რაც უაღრესად საჭიროა, რადგან, მას შემდეგ, რაც სამართალმცოდნეები გადაწყვეტენ, რომ წინადადების ესა თუ ის სტრუქტურა უფრო ზუსტად გადმოსცემს აზრს, **ლინგვისტური ცვლილების ან შესწორების შეტანა** თითქმის შეუძლებელია (Pescatore, 1999: 93).

**2.4. ევროკავშირთან ასოცირების ხელშეკრულების ინგლისურ-სლოვენური კორპუსი** - ასოცირების ხელშეკრულების თარგმნისას შექმნილ პრობლემებზე საუბრისას არაერთი ქვეყნის გახსენება შეიძლება; ყველაზე უპრიანია ისევ მცირე არეალზე გავრცელებული ენების მქონე ქვეყნების გაანალიზება, სადაც ხაზგასმულია ზუსტი ეკვივალენტობის საკითხი და მისი გადაჭრის გზები. ერთ-ერთი ასეთი მაგალითია სლოვენია.



სლოვენურ-ინგლისური პარალელური კორპუსი ოფიციალურად მოიცავს 15 დედანს და მათ თარგმანებს, რაც საერთო ჯამში შეადგენს 1 მილიონ სიტყვას. მათი ციფრული დარგობრივი ტერმინოლოგია კორპუსული მეთოდოლოგიითაა დალაგებული და სხვადასხვა ქვეკორპუსისგან შედგება. კორპუსის სტრუქტურაში ყველაზე დიდი (33%) ტერმინოლოგიური ბაზა მოიცავს იურიდიულ-სამართლებრივ და ისეთი სახის დოკუმენტებს, რომლებიც სლოვენის ევროკავშირთან ასოცირებას უკავშირდება.

სლოვენიური კორპუსის ანალიზისას მეცნიერები მივიდნენ დასკვნამდე, რომ საჭირო იყო შექმნილიყო ცალკე კორპუსი ასოცირების ხელშეკრულებაში არსებული ფრაზების უშუალოდ კონტექსტური მნიშვნელობებისთვის. რას ნიშნავს ეს ჩვენი დისერტაციის ფარგლებში? ფაქტია, რომ ტერმინისა და ფრაზის მნიშვნელობისადმი სლოვენიელი მეცნიერების მიდგომა ადასტურებს საკითხის ჩვენიული ხედვის სისწორეს, რადგან მათ, ირიბად მაგრამ მაინც, სწორედ ტერმინისა და ფრაზის ტელეოლოგიური მნიშვნელობით ცალკე გამოტანაზე, ახალ გაგებაზე გაამახვილეს ყურადღება, როცა ცალკე კორპუსის შექმნის საჭიროების საკითხი წამოსწიეს წინ და დააყენეს საკითხი, რომ ეს კორპუსი შედგენილიყო ამა თუ იმ ტერმინისა თუ ფრაზის არა სალექსიკონო, არამედ იმ კონტექსტური მნიშვნელობებით, რაც მათ უშუალოდ ასოცირების ხელშეკრულებაში შეიძინეს; ამ მნიშვნელობებს კი მხოლოდ მათი ტელეოლოგიური გაანალიზებითა და გაგებით თუ ჩავწვდებით, ანუ სამართლებრივი იურიდიული ტერმინებისა და ფრაზების უშუალოდ ასოცირების ხელშეკრულებაში შეძენილი კონტექსტური მნიშვნელობით. ამის შედეგად სამუშაო გაუადვილდებათ როგორც ლინგვისტებს, ისე ტერმინოლოგებს და მთარგმნელებს. ამრიგად, ფრა-

ზების, ტერმინოლოგიური ელექტრონული ბაზების შექმნა, პირველ რიგში, ხელსაყრელია მათთვის.

**თავი III. ტელეოლოგიური მიდგომები, როგორც სამართლებრივი ტექსტების ინტერპრეტაციის მეთოდი - შედგება სამი თავისაგან:**

**3.1. ტელეოლოგიის რაობა; ტექსტის ტელეოლოგიური ინტერპრეტაცია** - ტელეოლოგია წარმოდგება ბერძნული სიტყვისაგან „telos” (ბოლო, დაბოლოება). ის ჯერ კიდევ არისტოტელეს კვლევებიდან იღებს სათავეს, რომელიც ყოველი ორგანიზმის, არსების განვითარებას, სამყაროში არსებობას მისი გარკვეული დანიშნულების თვალსაზრისიდან გამომდინარე ხსნიდა. ყოველ სიტყვას, მსგავსად ცოცხალი ორგანიზმებისა, აქვს გარკვეული „დაბოლოება“ ანუ დანიშნულება, ფუნქცია. მკვლევართა ნაწილი კი ტერმინ „ტელეოლოგიას“ ფილოსოფოს ქრისტიან ვულფის სახელს უკავშირებს. ეს უკანასკნელი განიხილავდა მას, როგორც „ფილოსოფიის ნაწილს, რომელიც ხსნის საგანთა დაბოლოებებს ანუ მიზანს (დანიშნულებას)” (Glaserfeld, 2009:17). ტელეოლოგია განიხილავს და ეყრდნობა იმ საუკეთესოს, უმაღლეს ფასეულობას, სრულყოფილებას, რომლისკენაც მიილტვის სამყარო. სწორედ ეს გახლავთ სამყაროს არსებობის უმთავრესი მიზანი: სამყარო და მასში არსებული ყოველივე მატერიალური თუ არამატერიალური განუწყვეტლივ ვითარდება და იცვლება, მიილტვის უკეთესობისკენ.

მეცხრამეტე საუკუნის მეორე ნახევრიდან მოყოლებული ფილოსოფოსთა უმრავლესობა აღიარებდა საბოლოო მიზნის არსებობას და თავიანთ შრომებში განიხილავდნენ ტელეოლოგიური ანალიზის უპირატესობას (ვეველი, ჰერშელი და სხვანი). მიზანმიმართული ქცევებისა და საბოლოო მიზნის შესწავლა

მათი კვლევების ქვაკუთხედად იქცა, რითაც ხსნიდნენ ადამიანთა და სხვა ცოცხალ ორგანიზმთა (და არა მარტო ცოცხალ) მიზანმიმართულ ქცევას.

ასოცირების ხელშეკრულების პარალელურ რეჟიმში კვლევის პროცესში განსაკუთრებით საინტერესო და მნიშვნელოვანი გახლდათ ტელეოლოგიური ანალიზი, რომლის შედეგადაც დავრწმუნდით, რომ ყოველი სიტყვის არსებობა გარკვეული დანიშნულებით აიხსნება. არაერთხელ აღვნიშნეთ, რომ თარგმანში გამოყენებულ თითოეულ სიტყვას, მსგავსად ნებისმიერი ცოცხალი ორგანიზმისა, უნდა ჰქონდეს გარკვეული ფუნქცია. ტელეოლოგიური თვალსაზრისიდან გამომდინარე, იურიდიული დოკუმენტაციის ერთ-ერთი უმთავრესი თავისებურება გახლავთ ის, რომ თითოეულ ტერმინში იმთავითვე კოდირებული უნდა იყოს გარკვეული ინფორმაცია და, შესაბამისად, მას გააჩნდეს კონკრეტული მიზანი და დანიშნულება. აღნიშნულის გათვალისწინებით, მთარგმნელის უპირველესი ამოცანა უნდა იყოს იურიდიული ტერმინოლოგიის არა უბრალოდ თარგმნა, არამედ სამიზნე ტექსტისთვის ისეთი სიტყვების შერჩევა, რომლებშიც დედანში გამოყენებული სიტყვების მსგავსი ინფორმაციაა კოდირებული. უფრო მეტიც, მათ უნდა ჰქონდეთ იგივე მიზნები და დანიშნულება. ჩვენი კვლევის ფარგლებში პროგრამული ანალიზატორის მიერ გამოვლენილი მთარგმნელობითი უზუსტობების სიმრავლეს თუ გავითვალისწინებთ, ვრწმუნდებით, რომ იურიდიული ტერმინების დიდი ნაწილი თარგმნილია ლექსიკური მნიშვნელობით, უგულებელყოფილია მასში კოდირებული მიზანი.

მოსამართლემ უნდა მოახდინოს კანონის ინტერპრეტაცია, საამისოდ ის ვერ დაეყრდნობა მასში შემავალი ცალკეულ სიტყვებს, რომელთაც შესაძლოა სხვადასხვა მნიშვნელობა ჰქონდეს

და აზრობრივი ბუნდოვანება გამოიწვიოს. საუკეთესო გამოსავალია ტელეოლოგიური ინტერპრეტაცია, ერთგვარი ფილოსოფიური მიდგომა, რომელიც ეფუძნება ამათუიმ კანონის მიზნებს და მიმართულებებს და არა ცალკეული სიტყვებისა თუ ფრაზების მნიშვნელობას. ტელეოლოგიური ინტერპრეტაცია შეიძლება განვმარტოთ, როგორც „სასამართლოს მიერ გამოყენებული მეთოდი, როდესაც ისინი ახდენენ საკანონმდებლო აქტების ინტერპრეტაციას დასახული მიზნების, ღირებულებების, სამართლებრივი, სოციალური და ეკონომიკური ამოცანების ჭრილში.“

კანონის ტელეოლოგიური ინტერპრეტაციის უკეთ წარმოჩენის მიზნით, მიზანშეწონილად მიგვაჩნია განვიხილოთ რ.იაკობსონის ინტერპრეტაციის თეორია (ბერიძე, 2018: 16). ის განასხვავებს ლინგვისტური ნიშნის სამი ტიპის ინტერპრეტაციას, რომელთაგან კანონის ინტერპრეტაციასთან მიმართებით გამოვყოფდით **შიდაენობრივ თარგმანს** ანუ ვერბალური ნიშნების ინტერპრეტაციას იგივე ენაზე სხვა ნიშნებით.

კანონების შედგენა საკანონმდებლო ორგანოების პრეროგატივაა, მოსამართლეს კი მისი ინტერპრეტაციის სრული უფლება აქვს. ტელეოლოგიური ინტერპრეტაციით, მან კანონის გაგება და კონკრეტული განსახილველი საქმე ერთმანეთს უნდა მოარგოს. კანონის ინტერპრეტაციას ის ახდენს მშობლიურ ენაზე, ე.ი. ადგილი აქვს **შიდაენობრივ თარგმანს**.

საერთაშორისო დონის სამართლებრივი აქტების შემთხვევაში, როგორც ჩვენს მიერ გაანალიზებული ასოცირების ხელშეკრულებაა, არანაკლები მნიშვნელობა ენიჭება ინტერლიგვისტურ ინტერპრეტაციას, ანუ საკუთრივ თარგმანს. აუცილებელია, დაცული იყოს ეკვივალენტობა დედანსა და წყარო ტექსტს შორის, რათა მოსამართლემ სწორად განჭვრიტოს მასში კოდირებული მიზნები და მოარგოს კონკრეტულ განსახილველ

საქმეს, ანუ განახორციელოს კანონის ტელეოლოგიური ინტერპრეტაცია. იურიდიული პარალელური ტექსტების ტელეოლოგიური კვლევისას ვიხელმძღვანელებთ მოსაზრებით, რომ თითოეულ ტერმინში კოდირებული კონკრეტული მიზანი ემსახურება ადამიანის ცნობიერების მიმართვას კანონმორჩილები-საკენ, ხოლო მისი საბოლოო მიზანია სამართლიანი საზოგადოების ფორმირება. შესაბამისად, ტელეოლოგიური თვალსაზრისიდან გამომდინარე თუ ვიმსჯელებთ, შეგვიძლია დავასკვნათ: კანონები იქმნება არა იმისთვის, რომ უბრალოდ ვიცოდეთ, არამედ უნდა დავიცვათ, აღვასრულოთ ისინი, რათა ვიცხოვროთ კანონმორჩილ, სამართლიან საზოგადოებაში, რაც სოციუმის უზენაესი მიზანია. ყოველივე ზემოაღნიშნულიდან გამომდინარე, თარგმანში დაშვებული თითოეული შეცდომა ან უზუსტობა, ამა თუ იმ ტერმინის შეუსაბამო ეკვივალენტის შერჩევა, არ არის უბრალოდ მთარგმნელობითი შეცდომა, არამედ კანონის არასწორი ინტერპრეტაციის საფრთხე იკვეთება. მითუმეტეს, როდესაც საუბარია ისეთი დონის სახელმწიფოთაშორის იურიდიულ დოკუმენტზე, როგორც ასოცირების ხელშეკრულების თარგმანია. აღნიშნული ხელშეკრულების საფუძველზე საქართველო ერთიანი ევროპული სამართლებრივი სივრცის ნაწილი უნდა გახდეს, ამიტომ ქართულ თარგმანში გადმოტანილ თითოეულ იურიდიულ ტერმინში კოდირებული მიზანი დედანში არსებულ ტერმინში კოდირებულ მიზანს უნდა შეესაბამებოდეს.

**3.2. ტექსტის პრაგმატიკული ინტერპრეტაციის თავისებურებები** - ვინაიდან ტელეოლოგია მჭიდროდ უკავშირდება პრაგმატიკას, აღნიშნულ ქვეთავში მიზანშეწონილად მივიჩნიეთ ტექსტის პრაგმატიკული ინტერპრეტაციის თავისებურებების

განალიზება. პრაგმატიკულ მნიშვნელობაზე საუბრისას განვიხილეთ სხვადასხვა მკვლევრის მოსაზრება თვით ტერმინ პრაგმატიკის თაობაზე. გავრცელებულია მოსაზრება, რომ მისი თანამედროვე გამოყენება უკავშირდება ფილოსოფოს ჩარლზ მორისს (Levinson, 1983:2). მან პრაგმატიკა გამოყო სემიოტიკის (მეცნიერება ნიშნების შესახებ) ერთ-ერთ მიმართულებად, როგორც ნიშნების ინტერპრეტატორებთან მიმართების კვლევა.

მორისისეული გააზრებით, პრაგმატიკა არის ძალიან ფართო ცნება, რომელიც იკვლევს ნიშნების ფუნქციონირებისას მიმდინარე ფსიქოლოგიურ, ბიოლოგიურსა და სოციალურ პროცესებს. პრაგმატიკის ასეთი გაგება ძალიან ფართოა და ამ პოზიციის გაზიარების შემთხვევაში იგი მოიცავდა ენის ისეთ ასპექტებს, რომლებსაც ახლა იკვლევენ ლინგვისტიკის დამოუკიდებელი დისციპლინები: ფსიქოლინგვისტიკა, სოციოლინგვისტიკა და სხვა. მეორე მხრივ, პრაგმატიკის ჩარჩოები დავიწროვდა, რაც ფილოსოფოს კარნაპის სახელს უკავშირდება (Levinson, 1983:2). მისი აზრით, თუკი გამონათქვამის მნიშვნელობის კვლევისას მოსაუბრეს ღიად მოხსენიებენ, ასეთ შემთხვევაში ანალიზი მიმდინარეობს პრაგმატიკის ჭრილში. პოზიციას, რომლის მიხედვითაც პრაგმატიკა არის ენის იმ ასპექტის კვლევა, რომლებიც აუცილებლად მოითხოვენ ენის მომხმარებლის რეფერენციას, მივყავართ ტერმინის მეტად ვიწრო გაგებასთან. ლევისონის მტკიცებით (Levinson, 1983:7), პრაგმატიკის ამგვარი განმარტება შეუძლებელს გახდიდა ენის გამოყენების სრულყოფილ კვლევას. იგი ამტკიცებს, თუკი ლინგვისტიკისათვის მნიშვნელოვანია ტრიქოტომიის (სინტაქსი, სემანტიკა, პრაგმატიკა) არსებობა, მაშინ აუცილებელია პრაგმატიკისათვის სათანადო ჩარჩოს დადგენა, რათა ლინგვისტურმა კვლევამ მოიცვას კონტექსტის არსებითი ასპექტები. კონტექსტმა უნდა მოიცვას

დროისა და სივრცის პარამეტრები კომუნიკაციაში მონაწილე ინდივიდთა პიროვნებები, მათი რწმენა და მისწრაფებები, ცოდნა, განათლება და ინდივიდუალური ზრახვები. ამდენად, მნიშვნელოვანია მოძიება პრაგმატიკის ისეთი განმარტებისა, რომელიც ამ დისციპლინისათვის კვლევის სათანადო ჩარჩოებს დაადგენს და შესაძლებელს გახდის მნიშვნელობის სრულ და თანმიმდევრულ შესწავლას.

იულის (Yule, 2000:12) მიერ მოცემული ერთ-ერთი განმარტების შესაბამისად, პრაგმატიკა არის მოსაუბრის ან მწერლის მიერ გადმოცემული და მსმენელის ან მკითხველის მიერ ინტერპირირებული მნიშვნელობის კვლევა. შესაბამისად, პრაგმატიკა მოიაზრება, როგორც ადამიანის, ინდივიდის მიერ გამოხატულ გამონათქვამში ჩადებული ზრახვის ანალიზი და არა თავად ამ გამონათქვამში შემავალი სიტყვებისა თუ ფრაზების ანალიზი. ლევისონის მიერ შემოთავაზებული პრაგმატიკის მომდევნო განმარტების მიხედვით, ეს დისციპლინა არის ენის კვლევა ფუნქციური თვალსაზრისით და იგი არის მცდელობა, ახსნას ლინგვისტური სტრუქტურების კავშირი ექსტრა-ლინგვისტურ მიზეზებსა და ფაქტორებთან. ლევისონივე გვთავაზობს შემდეგ განმარტებას: პრაგმატიკა არის ენასა და კონტექსტს შორის არსებული იმ ურთიერთმიმართების კვლევა, რომელიც გრამატიზებულია ან კოდირებულია ენის სტრუქტურაში.

არსებობს გარკვეული მოსალოდნელი ან შესაფერისი გარემოებები, რომლებსაც წარმატებული კომუნიკაციის პირობებს უწოდებენ. შესაფერისი ანუ წარმატებული კომუნიკაციის ეს პირობები აუცილებელია, რათა კომუნიკაციაში მონაწილე ინდივიდებმა მოახდინონ შესრულებული სამეცხველო ქმედების სწორი და სათანადო ინტერპრეტაცია. ზოგიერთი ასეთი პირო-

ბათაგანი, რა თქმა უნდა, საერთოა ნებისმიერი სახის ენობრივი კომუნიკაციისათვის. მაგალითად, აუცილებელი წინაპირობაა, რომ მოსაუბრესა და მსმენელს ესმოდეთ ერთმანეთის, მაგალითად, უნდა საუბრობდნენ ერთსა და იმავე ენაზე. ასევე აუცილებელია ორივე ინდივიდს ჰქონდეს სმენის უნარი და ა. შ. ზოგიერთ შემთხვევაში სამეტყველო ქმედება იქნება წარუმატებელი, თუკი ქმედება სათანადო (შესაფერისი) პირის მიერ სათანადო პირობებში არ არის წარმოთქმული. მაგალითად:

–You are adjudged guilty! (თქვენ დამნაშავედ გცნობთ!)

თუკი ეს გამონათქვამი მოსამართლის მიერ სათანადო გარემოში არ არის წარმოთქმული, ამ შემთხვევაში სასამართლო პროცესის დროს სასამართლო დარბაზში, გამონათქვამი, რა თქმა უნდა, გადმოსცემს ინფორმაციას, მაგრამ იგი მთლიანად მოკლებულია ილოკუციურ ძალას.

**3.3. კონტექსტი და თარგმნის პრობლემა** - კომუნიკაციის მიმართ პრაგმატიკული მოსაზრებების მიხედვით, კონტექსტი მუდმივად განიცვლება, თუმცა ემპირიულად იგი დელიმიტირებულია (აქვს საზღვრები). სტალკნერის აზრით (Stalnaker, 1999:27), მოსაუბრის მიერ წარმოთქმული ყოველი პრესუპოზიცია კონტექსტის, როგორც ფონური ინფორმაციის, დელიმიტირებაა. თუმცა სპერბერისა და უილსონის (Sperber, Wilson, 1986:57) აზრით, კონტექსტი დელიმიტირებულია იმ ემპირიული ფაქტით, რომ მოსაუბრის გონებაში ზოგიერთი მოსაზრება აქტიურდება სხვა დამატებითი მოსაზრებებისა და ასოციაციების გააქტიურებით. ამ პოზიციის გაზიარება ნიშნავს იმის აღიარებას, რომ რაც უფრო მეტი დეტალი შემოდის კონტექსტში, იზრდება სამეტყველო ქმედების უკეთ გაგებისა და გააზრების შესაძლებლობა. ეს კი გულისხმობს, როგორც ზემოთ აღინიშნა,



სამეტყველო ქმედების მთელი მსოფლიო გამოცდილების შესაბამისად გაანალიზებას, რაც, რა თქმა უნდა, შეუძლებლად მიგვაჩნია. მნიშვნელობის კონტექსტზე დამოკიდებულებას ენის ფილოსოფოსები განიხილავენ ინდექსური, დემონსტრაციული, დესკრიპტული სიტყვებისა და მთელი პროპოზიციების ანალიზის პირობებში. კონტექსტზე დამოკიდებულებას უნდა ჰქონდეს რაიმე წყარო, რაც ნათლად ჩანს იმ ფაქტში, რომ ენის გამოყენება უმეტესად არის სიტუაციური. ფაქტობრივად, ენის გამოყენებელი პასუხისმგებელია, ენა გამოიყენოს რომელიმე ქმედების ჩარჩოში. სამეტყველო ქმედების ამგვარი სიტუაციურობა ლიმიტირებულ კონტექსტს მოითხოვს. თუკი სამეტყველო ქმედების რომელიმე კონტექსტის შესაბამისად შესრულება და გააზრება მას სიტუაციურ მოვლენად წარმოადგენს, მიგვაჩნია, რომ ეს სამეტყველო ქმედება ამავე კონტექსტის მიხედვით უნდა შეფასდეს. უფრო მეტიც, იმისათვის, რომ სამეტყველო ქმედება შეფასდეს, თავად კონტექსტი უნდა იყოს დელიმიტირებული, ანუ მკაფიოდ უნდა იყოს განსაზღვრული, თუ რა განეკუთვნება კონტექსტს და რა – არა. ყოვლისმომცველი კონტექსტი ვერ უზრუნველყოფს სიტუაციურობას და, როგორც ჯ. ოსტინი განმარტავს, იგი სამეტყველო ქმედებას აქცევს „იდეალურ გამონათქვამად, რომელიც შესაფერისი იქნებოდა ყოველი სიტუაციისათვის, ნებისმიერი მიზნისათვის“.

**თავი IV. იურიდიული პარალელური ტექსტების ანოტირება, სტრუქტურირება, თეგირება და კორპუსული კვლევა - შედგება ორი ქვეთავისაგან:**

**4.1. ინგლისური და ქართული იურიდიული პარალელური ტექსტების ანოტირების, სტრუქტურირებისა და თეგირების მოდელირება - მონაცემების ტექნოლოგიური ბაზა, რომელიც**

შეიქმნა კვლევის პროცესში, როგორც ადაპტირებული ღია პლატფორმა სპეციალიზებული ენობრივი კორპუსებისათვის. ის უზრუნველყოფს ქართული ენის პარალელური კორპუსის განვითარებას ზოგადად. ასევე, ის არის დაინტერესებული სპეციალისტებისათვის ერთიანი სამუშაო პლატფორმა, რომელიც აღჭურვილია კვლევის ჩვენეული მოდელით.

ნაშრომში გამოყენებულია კვლევის სპეციალიზებული სა-ძიებო კორპუსული პროგრამა, რომლის მეშვეობით ტერმინთა მოძიება, კონტექსტების გახსნა, კონტექსტური მნიშვნელობების დადგენა, დეფინიციების შეპირისპირებითი ანალიზი დროის მოკლე მონაკვეთში ხდება შესაძლებელი.

მსგავსად ნებისმიერი კვლევისა, კორპუსული კვლევის დროს უპირველესი და გადამწყვეტი მნიშვნელობა ენიჭება კვლევის მეთოდებს. გამომდინარე იქიდან, რომ კორპუსი ეყრდნობა დიდ მონაცემებს, (Big Data), მათი დამუშავება შეუძლებელია კვლევის ტრადიციული მეთოდების გამოყენებით, ეს აღემატება მრავალი მკვლევრის ერთობლივ ინტელექტუალურ შესაძლებლობებსაც კი. აუცილებელი ხდება სპეციალური კომპიუტერული პროგრამების გამოყენება. წინამდებარე ნაშრომში მონაცემთა ანალიზის სხვადასხვა მეთოდიდან უპირველეს ყოვლისა ვიყენებთ მონაცემთა მოპოვების (**Data Mining**) მეთოდს. მონაცემთა მოპოვება საკმაოდ ხანგრძლივი და შრომატევადი პროცესი გახლდათ, რომელიც განვახორციელეთ კორპუსლინგვისტიკაში აპრობირებულ სხვადასხვა ეტაპად: თავისუფალი ძიება (**Discovery**), რომლის დროსაც შევეცადეთ გარკვეული კანონზომიერებების გამოვლენას დედუქციური მეთოდით – კერძოდან ზოგადისაკენ; **ვალიდაცია**–იმის დასადასტურებლად, შეგვექმნა სწორი პროდუქტი; **ვერიფიცირება** – ამ ეტაპზე დავადასტურეთ, რომ შევქმენით ისეთი პროდუქტი, როგორსაც ვგეგ-

მავდით; პროგნოსტიკური მოდელირება (**Predictive Modeling**) – აღნიშნულ ეტაპზე გამოვლენილი კანონზომიერებები გამოვიყენეთ უცნობი მნიშვნელობის პროგნოზირებისათვის; გამონაკლისების ანალიზი (**Forensic Analysis**) – აღნიშნულ ეტაპზე კანონზომიერებებში მიკვლეული ანომალიები გამოვავლინეთ და ავხსენით.

მონაცემთა მოპოვების შემდეგი, არანაკლებ შრომატევადი და ხანგრძლივი პროცესი გახლდათ ტექსტის კოდირება, ანუ მოპოვებულ მონაცემთა სტრუქტურირება, რისთვისაც გამოვიყენეთ მარკირება – კორპუსში მონაცემთა სტრუქტურირების ყველაზე გავრცელებული საშუალება. როგორც უკვე აღვნიშნეთ, **მარკირება** (mark-up) შეიძლება განვმარტოთ, როგორც **მონიშვნა**, შესაბამისად, მარკირებისათვის გამოვიყენეთ თეგები, რომლებიც დავურთეთ სიტყვებს და ფრაზებს. შევიმუშავეთ თეგების (tag) სხვადასხვა ფორმა, რომლებიც ვიზუალურად შეიცავს ლათინურ ასო-ბგერებსა და გარშემორტყმულ სიმბოლოებს კუთხური ფრჩხილების სახით (<...>). ისინი წყვილებადაა წარმოდგენილი, მაგალითად: <lg> და </lg>. პირველი განიხილება, როგორც საწყისი თეგი, ხოლო მეორე – როგორც სასრული. საწყისი და სასრულ თეგებს შორის მოთავსებულია ტექსტი–საანალიზო მასალა, ანუ შიგთავსი. შიგთავსი თეგებში პატარა ასო-ბგერებითაა ჩაწერილი, მაგალითად: <lg>. ჩვენ მიერ შემუშავებული თეგები იერარქიული სტრუქტურისაა და ტექსტშია ჩაშენებული, რომელთა გამოყენებასაც დავაფუძნეთ საანალიზო ტექსტების მარკირება წინასწარ შექმნილი პროგრამული ანალიზატორისთვის.

როგორც აღვნიშნეთ, თეგირება დროში საკმაოდ გაწეილი, შრომატევადი და, იმავდროულად, საინტერესო პროცესია. იგი რამდენიმე ეტაპად განვავხორციელებთ.

ზემოთქმული მოსაზრებების გათვალისწინებით, თეგირების პირველ ეტაპზე შევეცადეთ სქემისა და მოდელის შერჩევას დასმული ამოცანის შესაბამისად, ანუ იმის გათვალისწინებით, თუ რა სახის ანალიზი გვსურდა, კერძოდ, ესაა ჩვენ მიერ შერჩეული საკვლევი თემის შემთხვევაში, თარგმნის ეკვივალენტობის პრობლემა პარალელურ ტექსტებში. თეგების ნაკრები შევადგინეთ იმის გათვალისწინებით, თუ რამდენად შეესაბამება ის როგორც ტექსტის თემას, ასევე დასახულ კვლევით სამუშაოს. მაგალითად, ჩვენს შემთხვევაში, კორპუსში შესატანი მასალის ავტომატურ და ნახევრად ავტომატურ რეჟიმში დამუშავების დროს თეგირება განვხორციელეთ რამდენიმე ეტაპად: 1. მორფოლოგიური ანალიზი; 2. სინტაქსური ანალიზი; 3, სემანტიკური ანალიზი; 4. ტელეოლოგიური ანალიზი. ქვემოთ შემოგთავაზებთ კონკრეტულ მაგალითებს, მათ ანალიზსა და დასკვნებს.

ჩვენ მიერ ჩატარებულ კვლევაში უპირატესობას ვანიჭებთ ვიზუალიზაციის ისეთ ფორმებს, როგორებიცაა გრაფები და ცხრილები. ისინი ფართოდ გამოიყენება კორპუსის აგებისთვის საჭირო სტრუქტურირებული მონაცემების აღსაწერად, ასევე სიტყვების/თეგების ღრუბელი (word/tag cloud), რომელიც დღესდღეობით ტექსტური მონაცემების ვიზუალიზაციის საკმაოდ გავრცელებული ფორმაა.

### **კორპუსის ტექნიკური ფორმატი**

კორპუსი დაშიფრულია XML ფორმატში, ტექსტის დაშიფვრის ძირითადი რეკომენდაციების (TEI P4) შესაბამისად. კორპუსი შედგება ორი ნაწილისგან: ტექსტებისა და განმარტებებისაგან.

ტექსტები დაჯგუფებულია ენების მიხედვით (ინგლისური და ქართული). ერთ ენაზე არსებული ტექსტები ქმნიან ერთ

TEI კორპუსს, რომელიც შედგება საკუთრივ ტექსტისა და TEI სათაურისაგან (ე.წ. ჰედერი). ეს უკანასკნელი გვაწვდის ამომწურავ ინფორმაციას კორპუსის შესახებ. ყოველი ტექსტი თავის მხრივ შეიცავს TEI სათაურსა და საკუთრივ ტექსტს, რომელიც ცალკეული აბზაცისაგან შედგება.

TEI ტექსტს თითოეული ენობრივი წყვილისთვის თან ერთვის ორმაგი განმარტება, თუმცაღა ტექსტი სრულად არ არის წარმოდგენილი, არამედ ცალკეული აბზაცის ამონარიდებს შეიცავს.

კვლევისას გამოვიყენეთ ჩვენი უცხოელი კოლეგის მიერ გაზიარებული გამოცდილება, კერძოდ: 2019 წლის ოქტომბერში ბათუმის შოთა რუსთაველის სახელმწიფო უნივერსიტეტის ჰუმანიტარულ მეცნიერებათა ფაკულტეტთან არსებულ თარგმანისა და ინტერდისციპლინარული კვლევების ცენტრს ეწვია საარლენდის (გერმანია) უნივერსიტეტის პროფესორი ეკატერინა ლაფშინოვა-კოლტუნსკი. პროფ. ლაფშინოვა-კოლტუნსკიმ პრეზენტაციის სახით წარმოგვიდგინა კორპუსის განმარტებები სხვადასხვა ავტორების მიერ. ერთმანეთს შევადარეთ ერთ-ერთოვანი და მულტილინგვური კორპუსები. ორენოვან და მრავალენოვან კორპუსებში ტექსტები დალაგებულია რეგისტრისა და ჟანრის მიხედვით.

პროფესორ ლაფშინოვას დახმარებით, ასოცირების ხელსეკრულების ქართული ტექსტის ფრაგმენტისა და რამდენიმე თეგის გამოყენებით, შევქმენით ნიმუში MMAX2 პროგრამაში ანოტაციისთვის, საქარალდე სახელწოდებით “project”. ვხსნით აღნიშნულ საქაღალდეს, ეკრანზე ჩნდება კიდევ ერთი საქაღალდე სახელწოდებით MMAX-terminology, რომელსაც ისეთივე აგებულება გააჩნია, როგორც ყველა დანარჩენ MMAX პროექტებს (როგორცაა: Basedata, Markabls, etc.)

ვხსნით პროგრამას MMAX2 და ვპოულობთ ზემოაღნიშნულ საქაღალდეს (MMAX-terminology) საძიებო ველის File>Load საშუალებით MMAX2 ფანჯარაში, გამოჩნდება ფაილი AA-geo.mmax, რომლის გახსნისას ვიხილავთ ასოცირების ხელსეკრულების ქართულენოვანი ტექსტის ნაწყვეტს XML ფორმატში და ანოტაციის პროგრამა დაიწყება. აღსანიშნავია, რომ MMAX2 პროგრამა ევროკავშირის წევრი თუ კანდიდატი ქვეყნების თითქმის ყველა ენას მოიცავს, გარდა ქართულისა. ეს გახლდათ მოცემულ პროგრამაში ქართულენოვანი ტექსტის შეყვანის პირველი წარმატებული მცდელობა.

#### **4.2. ანოტირებული, სტრუქტურირებული და თეგირებული პარალელური იურიდიული ტექსტების კორპუსული კვლევა ავტომატურ და ნახევრავტომატურ რეჟიმებში**

კვლევის პირველ ეტაპზე საქართველო-ევროკავშირის ასოცირების შეთანხმების (კარი 1-4, მუხლი 1-276) ინგლისურ ტექსტში და ქართულ თარგმანში თეგებით მოვნიშნეთ კოლოკაციები, კალკირებები, უცხოური სიტყვები, ტერმინები (იურიდიული და ეკონომიკური), სინტაქსური კალკირება, ვნებითი გვარის კონსტრუქციის არასწორი ინტერპრეტაცია, ტავტოლოგია, მართლწერა, მთარგმნელობითი შეცდომა, აბრევიატურა, ლექსიკური დამატება. იხილეთ მაგალითი:

The Parties <lng\_clc>commit themselves</lng\_clc> to the <lg\_trm>rule of law</lg\_trm>, <lng\_clc>good governance</lng\_clc>, the <lng\_clc>fight against corruption</lng\_clc>, the fight against the various forms of transnational <lg\_trm>organised crime and terrorism</lg\_trm>, the promotion of sustainable development, effective multilateralism and the fight against the proliferation of <lg\_trm>weapons of mass destruction</lg\_trm> and their <lng\_trm>delivery

systems</lng\_trm>. This commitment constitutes a <lng\_clc>key factor</lng\_clc> in the development of the relations and cooperation between the Parties and contributes to regional peace and stability.

მხარეები <lng\_clc>ვალდებულებას იღებენ</lng\_clc> უზრუნველყონ <lg\_trm>კანონის უზენაესობა</lg\_trm>, განახორციელონ <lng\_calq>კარგი მმართველობა</lng\_calq>, <lng\_clc>ებრძოლონ კორუფციას</lng\_clc> და სხვადასხვა ფორმის <soc\_fwd> ტრანსნაციონალურ</soc\_fwd> <lg\_trm>ორგანიზებულ დანაშაულსა და ტერორიზმს</lg\_trm>, ხელი შეუწყონ <lng\_calq> მდგრად განვითარებას</lng\_calq>, ეფექტიან <soc\_fwd>მულტილატერალიზმს</soc\_fwd> და <synt\_calq>იბრძოლონ <lg\_trm>მასობრივი განადგურების იარაღის</lg\_trm> და მისი <lng\_trm>მიწოდების საშუალებების</lng\_trm> გავრცელების წინააღმდეგ</synt\_calq>. ეს ვალდებულებები წარმოადგენს მხარეებს შორის ურთიერთობებისა და თანამშრომლობის განვითარების <lng\_calq><lng\_clc>ძირითად ფაქტორს</lng\_clc></lng\_calq> და ხელს უწყობს მშვიდობასა და სტაბილურობას რეგიონში.

ჩვენი სადისერტაციო ნაშრომის კვლევის მიზნებიდან გამომდინარე, სხვადასხვა ლინგვისტური კონსტრუქციისა და ტერმინის მოსანიშნად გამოვიყენეთ შემდეგი თეგები:

- <lng\_clc>...</lng\_clc> კოლოკაცია
- <lng\_calq>...</lng\_calq> კალკირება
- <soc\_fwd>...</soc\_fwd> უცხოური სიტყვები/ტრანსლიტერაცია
- <lng\_trm>...</lng\_trm> ტერმინები
- <lg\_trm>...</lg\_trm> იურიდიული ტერმინები
- <ecn\_trm>...</ecn\_trm> ეკონომიკური ტერმინი
- <synt\_calq>...</synt\_calq> სინტაქსური კალკირება

<lng\_mpv>...</lng\_mpv> ვნებითი გვარის კონსტრუქციის არასწორი

ინტერპრეტაცია

<st\_taut>...</st\_taut> ტავტოლოგია

<lng\_spl>...</lng\_spl> მართლწერა

<tr\_err>...</tr\_err> მთარგმნელობითი შეცდომა

<lng\_abr>...</lng\_abr> აბრევიატურა

<tr\_lex.add>...</tr\_lex.add> ლექსიკური დამატება

თეგები შემუშავებულია ბათუმის შოთა რუსთაველის სახელმწიფო უნივერსიტეტის ევროპეისტიკის დეპარტამენტის ასოცირებული პროფესორის ხათუნა ბერიძის მიერ. თეგირების პროცესი მოითხოვს განსაკუთრებულ სიფრთხილესა და ყურადღებას. ის რამდენიმე ეტაპად განვახორციელებთ: 1. ასოცირების ხელშეკრულების ინგლისური და ქართული ტექსტების პარალელურ რეჟიმში დამუშავება-შედარება, რასაც ძირითადად გაცნობითი ხასიათი ჰქონდა; 2. წყარო და სამიზნე ტექსტების მარკირება პარალელურ რეჟიმში, ერთი და იმავე თეგის გამოყენებით – მარკირების პირველი ეტაპი; 3. უკვე მარკირებული ტექსტების ხელმეორედ გადახედვა – მარკირება. უმეტესწილად აღნიშნული ეტაპის გამეორება რამდენჯერმე გვიხდებოდა, ვინაიდან მარკირებისას დაშვებული მცირედი უზუსტობაც კი პრობლემას გვიქმნიდა – პროგრამული ანალიზატორი ჩვენ მიერ მიწოდებულ მონაცემებს ვერ აღიქვამდა.

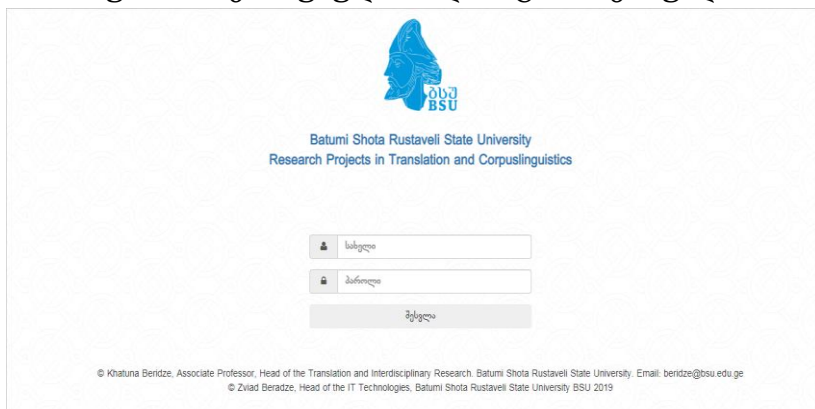
მარკირების ხანგრძლივი და შრომატევადი პროცესის შემდეგ ინგლისური და ქართული ტექსტები Word-ის ფორმატიდან გადავიყვანეთ XML ფორმატში. საამისოდ გამოვიყენეთ სპეციალური პროგრამა NOTEPAD++, რომლის უპირატესობა გახლავთ ის, რომ პროგრამული ანალიზატორისთვის აღქმადია, ამასთანავე, გვიადვილებს მუშაობას იმ შემთხვევაში, თუკი რომელიმე





მომდევნო ეტაპზე მარკირებული ტექსტების გადასამოწმებლად გამოვიყენეთ ბათუმის შოთა რუსთაველის სახელმწიფო უნივერსიტეტის თარგმანისა და ინტერდისციპლინური კვლევების ცენტრისა და ტექნოლოგიური დეპარტამენტის თანამშრომლობით შექმნილ სპეციალურ პროგრამა. იხილეთ სურ.3:

სურათი 3. კორპუსული ანალიზატორის ვიზუალი:



The image shows a web interface for a research project. At the top center is the logo of Batumi Shota Rustaveli State University (BSU), featuring a blue profile of a head. Below the logo, the text reads "Batumi Shota Rustaveli State University" and "Research Projects in Translation and Corpus Linguistics". There are three input fields: the first is labeled "სახელი" (Name) with a person icon, the second is labeled "პაროლი" (Password) with a lock icon, and the third is a "შეყვლა" (Login) button. At the bottom, there is a copyright notice: "© Khutuna Beridze, Associate Professor, Head of the Translation and Interdisciplinary Research, Batumi Shota Rustaveli State University, Email: beridze@bsu.edu.ge © Zviad Beradze, Head of the IT Technologies, Batumi Shota Rustaveli State University BSU 2019".

აღნიშნული პროგრამით სარგებლობისთვის თავდაპირველად საჭიროა დავრეგისტრირდეთ, როგორც მომხმარებელი. მომხმარებლის საკუთარი სახელისა და პაროლის მითითებით შევდივართ პროგრამაში. იხილეთ სურათი 4:

## სურათი 4. პროგრამის ვიზუალი

შეავარი
ტექსტის დამატება/რედაქტირება
ფაილის ავირთვა
გაძლია

აირჩიეთ თემა	<input type="text"/>	
კოლოკაცია	<input type="checkbox"/>	კალიგრაფია <input type="checkbox"/>
უცხოური სიტყვები	<input type="checkbox"/>	ნასესხობა <input type="checkbox"/>
ტერმინები	<input type="checkbox"/>	იურიდიული ტერმინები <input type="checkbox"/>
ეკონომიკური ტერმინი	<input type="checkbox"/>	თარგუმის ცრუ მგობრები <input type="checkbox"/>
სინტაქსური კალიგრაფია	<input type="checkbox"/>	ენებ ვეარის კონსტრ.არასწორი ინტერპრეტაცია <input type="checkbox"/>
ტავტოლოგია	<input type="checkbox"/>	მართლწერა <input type="checkbox"/>
მთარგმნელობითი შეუღობა	<input type="checkbox"/>	აბრევიატურა <input type="checkbox"/>
ლექსიკური დამატება	<input type="checkbox"/>	

უპირველეს ყოვლისა, პროგრამაში ავტვირთავთ XML ფორმატში გადაყვანილი ინგლისურსა და ქართულ ტექსტებს. შემდეგ სპეციალურ ველში –**თეგების დამატება** შეგვყავს ჩვენ მიერ წინასწარ შემუშავებული თეგები. ავირჩევთ XML ფორმატში უკვე შეყვანილ სასურველ ტექსტს, ველში **ავირჩევთ თემას**, მოვნიშნავთ სასურველ თეგებს და ვაჭერთ ღილაკს **გაგზავნა**. ყოველივე აღნიშნულის შემდეგ პროგრამა ავტომატურ რეჟიმში ეძებს და გვაწოდებს სტატისტიკურ მონაცემებს, რომლებიც შეგვიძლია გამოვიყენოთ ჩვენი კვლევის პროცესში დამუშავებული ორენოვანი მასალის შეპირისპირებითი ანალიზისთვის. პროგრამა ინგლისურსა და ქართულ ტექსტებს პარალელურად ალაგებს და, რაც ყველაზე მნიშვნელოვანია, რაოდენობრივი კვლევის მეთოდოლოგიისთვის ავტომატურ რეჟიმში გვამძღვეს თეგირებული ტერმინების სტატისტიკას. იხილეთ ეკრანის სურათი, რომელიც ასახავს პროგრამის მიერ პარალელურად დალაგებული ორენოვანი ტექსტის სტატისტიკური ანალიზატორიდან მოპოვებულ მონაცემებს:



რაც შეეხება ანალიზატორის შედეგებს:

გამოვავლინეთ და გავანალიზეთ მთარგმნელობითი უზუსტობებისა და შეცდომების მთელი რიგი შემთხვევები, სახელდობრ: კალკირება - 47 მაგალითი; სინტაქსური კალკირება - 18 შემთხვევა; უცხოური სიტყვების და ნასესხობების უმართებულო გამოყენება - 42 მაგალითი; ტერმინები - 230, რომელთა გამოყენებითაც შეიძლება საქართველოს ევროკავშირთან ასოცირების ტერმინოლოგიური ლექსიკონის შედგენა.

### **დასკვნა**

მოცემული ნაშრომის მიზანი გახლდათ იურიდიული ტექსტების თარგმანის ანალიზის მოდელირების პრეცედენტის შექმნა პარალელურ კორპუსში, რადგანაც კორპუსული კვლევა თანამედროვე თარგმანმცოდნეობაში განსაკუთრებით აქტუალურია. არჩევანი შევაჩერეთ ასოცირების ხელშეკრულებაზე, რომლის ორიგინალი და სამიზნე ტექსტების შეპირისპირებითი ანალიზის შედეგად მივედით შემდეგ დასკვნებამდე:

- შევქმენით იურიდიული ტექსტების თარგმანის ანალიზის ინდივიდუალური მოდელი ინგლისური და ქართული ტექსტების პარალელური კორპუსისთვის;

- ჩვენი კვლევის ფარგლებში შექმნილი თარგმანის კორპუსი მკვლევარს კონკრეტული კვლევის საკითხის მოკლე დროში გადაწყვეტის საშუალებას აძლევს, შეუძლია შეარჩიოს და გამოიყენოს კორპუსის ტექსტები სხვადასხვა ტიპის შეპირისპირებითი კვლევისათვის;

- აღნიშნული პარალელური კორპუსი ღირებული ენობრივი რესურსია, რომელიც თარგმნის პროცესს მდიდარი ლექსიკით ამარაგებს და მთარგმნელისთვის მზა ფრაზების ერთ-

გვარ ბანკს წარმოადგენს, ამარტივებს და აჩქარებს მის საქმიანობას;

- ნაშრომში გამოყენებული სპეციალიზებული საძიებო პროგრამის მეშვეობით ტერმინთა მოძიება, კონტექსტების გახსნა, კონტექსტური მნიშვნელობების დადგენა, დეფინიციების შეპირისპირებითი ანალიზი დროის მოკლე მონაკვეთში ხდება შესაძლებელი;

- კვლევის მიზნების შესაბამისად, პარალელური კორპუსის საშუალებით ავტომატურ რეჟიმში გამოვიკვლიეთ მთარგმნელობითი უზუსტობების მთელი რიგი კატეგორიები და მოკლე დროში მივიღეთ მათი სტატისტიკა;

- თუკი ადრეულ პერიოდში ლინგვისტური კვლევები ძირითადად თეორიით შემოიფარგლებოდა და შემდეგ ხდებოდა მათი პრაქტიკაში დანერგვა (პროცესი „ზემოდან - ქვემოთ“), ჩვენი კორპუსლინგვისტური ნაშრომი უშუალოდ პრაქტიკულ კვლევებს ეყრდნობა (პროცესი „ქვემოდან - ზევით“). უფრო მეტიც, ჩვენს მიერ თეორიულად დამტკიცებული მასალა კვლავ პრაქტიკული გამოყენებისთვისაა განკუთვნილი;

- ჩვენი კვლევის ფარგლებში პროგრამული ანალიზატორის მეშვეობით აღმოჩენილი მთარგმნელობითი უზუსტობებისა და შეცდომების გათვალისწინება ხელს შეუწყობს: ა) ევროკომისიის თარგმანის ხარისხის სტანდარტებთან მიახლოებას, რომლებიც ევროკომისიის თარგმანის გენერალური დირექტორატის მიერ არის ჩამოყალიბებული და რაც მომავალში, ასოცირების შემდეგ, გახდება სავალდებულო სახელმძღვანელო-გზამკვლევი საქართველოში თარგმანის ხარისხის სტანდარტიზაციისათვის; ბ) ასოცირების ხელშეკრულების ქართული თარგმანის მიახლოებას საერთაშორისო სტანდარტებით განსაზღვრულ ძირითად პრინციპებსა და თარგმანის ხარისხის მოთ-

ხოვნებთან, რადგანაც ევროკავშირის სამართლებრივი აქტების ყველა ენაზე არსებული ვერსია თანაბრად ავთენტური უნდა იყოს; გ) ხელშეკრულების ქართული ვერსიის კიდევ უფრო მიახლოებას ზოგადსაკანონმდებლო პრინციპებთან, როგორცაა: კანონის წინაშე თანასწორუფლებიანობა და კანონის სანდოობა; დ) თავიდან აგვაცილებს სამართალწარმოების მოსალოდნელ რისკებს, მათ შორის ფინანსურ, პოლიტიკურ თუ იმიჯთან დაკავშირებულ ზიანს; ე) შეამცირებს სირთულეებსა და პრობლემებს კანონის ინტერპრეტაციისას ან კანონის ეროვნულ დონეზე დანერგვისას;

- ჩვენი კვლევის ფარგლებში შექმნილი იურიდიული პარალელური კორპუსის პრაქტიკულ ღირებულებათაგან აღსანიშნავია შემდეგი: მისი გამოყენება შეიძლება ა) სტატისტიკური მანქანური თარგმანის, ბ)ორენოვანი ლექსიკური ან სემანტიკური რესურსების შესამუშავებლად; გ) ლექსიკონების ან ანთოლოგიებისათვის; დ) მრავალენოვანი ინფორმაციის მოძიების პროგრამული უზრუნველყოფის გამოცდისა და სწავლებისთვის;

- რაც ყველაზე მნიშვნელოვანია, აღნიშნული კვლევის ფარგლებში პროგრამული ანალიზატორის მიერ გამოკვეთილი ტერმინების თარგმნის მაგალითებით შეიძლება საქართველოს ევროკავშირთან ასოცირების ტერმინოლოგიური ლექსიკონის შედგენა;

- ჩვენს მიერ შექმნილი ასოცირების ხელშეკრულების ინგლისურ-ქართული პარალელური კორპუსის ყველაზე დიდ პრაქტიკულ ღირებულებად მიგვაჩნია ის ფაქტი, რომ ეს არის მცდელობა გაეზღვეთ დღესდღეობით არსებული ყველაზე მასშტაბური მრავალენოვანი კორპუსის **JRC-Acquis**-ის ნაწილი, რომელიც ევროკავშირის წევრი ქვეყნების ოცზე მეტ ენას მოიცავს;

- განსაკუთრებით აღნიშნავდით იმ ფაქტს, რომ ჩვენს მიერ შექმნილ იურიდიულ პარალელურ კორპუსს გააჩნია კიდევ ერთი წინგადადგმული ნაბიჯი, წინაპირობა JRC-Acquis-ში გაწევრიანებისათვის: ასოცირების ხელშეკრულების ინგლისური და ქართული თეგირებული ტექსტები უკვე გადაყვანილი გვაქვს XML ფორმატში, მაშინ როდესაც ევროკავშირის კანდიდატი ქვეყნების ენებზე არსებული დოკუმენტაცია დღესდღეობით მხოლოდ Word-ის ფორმატშია ხელმისაწვდომი, რასაც JRC-Acquis-ის პროგრამა ვერ აღიქვამს;

- აღნიშნული კვლევის ფარგლებში პროგრამული ანალიზატორის მიერ გამოვლენილი სტატისტიკის გათვალისწინებით შესაძლებელი გახდება თარგმანში დაშვებული შეცდომების ან უზუსტობების რაოდენობის შემცირება, რაც ზოგჯერ ძალიან საყურადღებოა, ვინაიდან ასოცირების ხელშეკრულების მსგავსი სახელმწიფოთაშორისი მნიშვნელობის მქონე იურიდიული დოკუმენტი, შეიძლება შეიცავდეს განსაკუთრებით საგულისხმო ინფორმაციას, მაგალითად, ატომური იარაღის გავრცელების შესახებ. მსგავს შემთხვევაში მთარგმნელობითი შეცდომით გამოწვეული სავალალო შედეგის რისკი შემცირდება;

- ჩვენი კვლევის კიდევ ერთი პრაქტიკული ღირებულებაა პროგრამული ანალიზატორის ფუნქცია პარალელურად დაალაგოს ორენოვანი ტექსტები და, რაც ყველაზე მნიშვნელოვანია, ავტომატურ რეჟიმში მოგვცეს თეგირებული ტერმინების სტატისტიკა;

- განსაკუთრებით აღსანიშნავია ის ფაქტი, რომ კვლევის შედეგების გადამოწმება -ვერიფიკაცია მოცემული პროგრამის საშუალებით შესაძლებელია დისტანციურად. შეგვიძლია თამამად განვაცხადოთ, რომ კორპუსლინგვისტიკისა და ინფორმა-



ტიკის ერთობლივი და უდიდესი შედეგი მიღწეულია - შეგვიძლია კვლევის შედეგის მიღება უმოკლეს დროში;

- მიგვაჩნია, რომ ჩვენი დისერტაციის ფარგლებში შექმნილი ქართულ-ინგლისური პარალელური კორპუსი არის ერთ-ერთი წინადადგმული ნაბიჯი და მცდელობა იმისა, რომ მსოფლიოში არსებულ მრავალენოვან კორპუსებს შორის ქართულენოვანმა კორპუსმაც დაიმკვიდროს ადგილი და ხელმისაწვდომი გახდეს დაინტერესებული სპეციალისტებისთვის.

- ანალიზატორით მიღებული სტატისტიკური მაჩვენებლების, პარალელურად განთავსებული ტერმინოლოგიური ლექსიკონისა და პარალელურად სტრუქტურირებული ტექსტების მიღებით, ჩვენ პროგრამული თვალსაზრისით მივალწიეთ ჩვენი ნაშრომის მიზანს და შევქმენით იურიდიული ტექსტების თარგმანის მანუალური მარკირების, პროგრამულ ანალიზატორში დამუშავებისა და შემდგომი სამეცნიერო კვლევის მოდელირების პრეცედენტი ინგლისური და ქართული პარალელური კორპუსებისთვის.

- ჩვენი კვლევის ფარგლებში განვსაზღვრეთ, რომ პროგრამული ანალიზატორის როლი და ფუნქცია უაღრესად მნიშვნელოვანია თარგმანის ხარისხობრივი კვლევისთვის, რადგან ავტომატურად მიღებული რაოდენობრივი მაჩვენებლები იძლევა კვლევის სისტემატიზირებულად წარმართვის და კონკრეტული კვლევის საკითხის მოკლე დროში გადაწყვეტის საშუალებას.