

Հակոբյան Տրիդա (Հայաստան,

ԳԱԱ լեզվի ինստիտուտ)

**ՀԱՅԵՐԵՆ ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԻ ՍՏԵՂԾՄԱՆ  
ՆՊԱՏԱԿԸ ԵՎ ԽՆԴԻՐՆԵՐԸ**

Էլեկտրոնային եղանակով տեքստերի կազմումը և դրանց հրապարակումն ու տարածումը համացանցի միջոցով գրեթե լիովին փոխարինել են ձեռագիր տեքստերին՝ աստիճանաբար կարևորելով տեքստի լեզվական որակի բարձրացման խնդիրը: Հետևաբար համակարգչային ծրագրերի՝ լեզվաբանության մեջ տեսական և գործնական տարբեր նպատակներով կիրառություններից ամենահիմնականը կարելի է համարել համակարգչային լեզվաբանների կողմից տեքստի ավտոմատ սրբագրման և խմբագրման համակարգերի ստեղծումը, որոնք ունեն ընդհանրական այնպիսի հատկանիշներ, ինչպիսիք են ուղղագրական, քերականական, քիչ դեպքերում նաև՝ կետադրական սխալների ուղղումը: Կատարելով գրեթե նույն գործառնությունը՝ սրբագրման բոլոր համակարգերի ու ծրագրերի տարբերակիչ ամենահիմնական հատկանիշն այն է, որ դրանք տեքստում եղած սխալներն ուղղումն են տարբեր կերպ և ոչ ամբողջությամբ՝ պայմանավորված համակարգի զարգացվածության աստիճանով: Օգտատերերի շրջանում առավել մեծ հեղինակություն են վայելում Grammarly<sup>1</sup>, Ginger<sup>2</sup>, SpellCheckPlus Pro<sup>3</sup>, CorrectEnglish Complete<sup>4</sup>, WhiteSmoke<sup>5</sup> օտարալեզու սրբագրման համակարգերը:

Հայերեն տեքստերի էլեկտրոնային սրբագրման փոքրաթիվ ծրագրեր ևս ստեղծվել են: Դրանցից իրենց տարածվածությամբ և համեմատաբար լայն կիրառությամբ առանձնանում են Armenian PowerSpell համակարգչային ծրագիրը<sup>6</sup> և HySpell.com կայքում գործող HySpell ծրագիրը<sup>7</sup>: Սրանց միջոցով հնարավոր է կատարել տեքստում եղած սխալների ոչ բավարար մակարդակով ուղղում, ուստի կարող ենք ասել, որ հայերեն տեքստերի սրբագրման համակարգերը որակական առումով շատ թերի են: Դա վերաբերում է ինչպես բառապաշարի ընդգրկման ծավալին, այնպես էլ քերականությանը, քանի որ քերականական ոչ բոլոր տեսակի սխալներն է հնարավոր ուղղել դրանց միջոցով:

<sup>1</sup> <https://www.grammarly.com/>

<sup>2</sup> <http://www.gingersoftware.com/grammarcheck>

<sup>3</sup> <https://pro.spellcheckplus.com/>

<sup>4</sup> <http://www.correctenglish.com/>

<sup>5</sup> <http://www.whitesmoke.com/>

<sup>6</sup> <http://www.powerspell.am/>

<sup>7</sup> <http://hyspell.com/>

Այս համատեքստում մեծապես կարևորում ենք էլեկտրոնային սրբագրման նոր՝ առավել լիակատար և ամբողջական համակարգի ստեղծման խնդիրը, ուստի սույն հոդվածում կանդրադառնանք հայերեն էլեկտրոնային սրբագրման համակարգի կազմության՝ մեր կողմից մշակված սկզբունքներին, որոնք վերաբերում են համակարգի ստեղծման նպատակին ու խնդիրներին, կառուցվածքին ու ընդգրկման ծավալին:

Հայալեզու սրբագրման նոր՝ առավել լիակատար համակարգի ստեղծման կարևորությունը պայմանավորված է հայերեն տեքստերի լեզվական որակի բարձրացման անհրաժեշտությամբ: Ընդունված է ասել, որ գրավոր խոսքը բանավորի համեմատությամբ առավել կանոնարկված է, սակայն դա ամեննիստ էլ չի բացառում ամենատարբեր սխալների առկայությունը, շատ դեպքերում նաև՝ առատությունը տեքստում: Ուստի, նորաստեղծ էլեկտրոնային սրբագրման համակարգի խնդիրը տեքստում եղած սխալների առավելագույն չափով ուղղումն է և, որպես արդյունք, հնարավորինս քիչ սխալներ պարունակող տեքստեր ներկայացնելը:

Ամենից առաջ հարկ է ընդգծել, որ նման բազմագործառույթ համակարգի ստեղծումը բարդ և ժամանակատար խնդիր է, որը պահանջում է համակարգչային ծրագրերի համար կիրառելի դարձնելու նպատակով լեզվի բառապաշարի, ինչպես նաև քերականական համակարգի մշակման, ձևայնացման ուղղությամբ լեզվաբանների և ծրագրավորողների համակողմանի աշխատանք: Այնուամենայնիվ, նմանատիպ գործառույթ իրականացնելու նպատակով ստեղծված այլ ծրագրերի համեմատությամբ այն կունենա մի շարք ընդգծված տարբերություններ, որոնք վերաբերում են ինչպես էլեկտրոնային համակարգի ձևային, այնպես էլ բովանդակային կողմին:

Քանի որ դրա ստեղծման առաջնային նպատակը հայախոս հանրության շրջանում գրավոր խոսքը որակական զարգացման նոր մակարդակի հասցնելն է, նպատակահարմար է՝ այն գործի համացանցային ինքնուրույն կայքում լիովին անվճար և առանց որևէ ծրագրի ներբեռնման անհրաժեշտության: Մինչդեռ հայերեն տեքստերի սրբագրման մյուս ծրագրերի օգտագործման հնարավորությունները շատ սահմանափակ են: Օրինակ՝ PowerSpell համակարգչային ծրագրից օգտվելու համար հարկ է գնել դրա ամբողջական փաթեթը, ավելացնել Microsoft Word, Excel, PowerPoint տեքստային խմբագրիչներին և օգտագործել համակարգչում: Մա ենթադրում է, որ ոչ միայն սահմանափակ թվով մարդիկ միայն կարող են օգտվել դրանցից, այլև ծրագրի ստեղծողներն անգամ չեն կարող այն հաճախակի փոփոխել, համալրել կամ ուղղել ի հայտ եկած սխալները, քանի որ դա անելու միակ եղանակը ծրագրի նոր՝ լրամշակված տարբերակի թողարկումն է: Առանձին կայ-

քում գործելու շնորհիվ՝ էլեկտրոնային սրբագրման նոր համակարգը, սակայն, կայքի հեղինակներին այն մշտապես թարմացնելու, փոփոխելու, նկատված սխալները կամ ծրագրային թերացումները անհապաղ շտկելու, առավել զարգացնելու ու լիակատար դարձնելու հնարավորություն կտա:

Հայերեն տեքստերի սրբագրման համեմատաբար հաջող ծրագրերից երկրորդը HySpell.com կայքում գործող HySpell համակարգչային ծրագիրն է և համապատասխան ծրագրային հավելվածը: Որպես սրբագրման ծրագիր՝ այն ունի համակարգչային նախորդ ծրագրին բնորոշ բոլոր թերությունները: HySpell ծրագրային հավելվածի հասանելիությունը ևս սահմանափակվում է մի շարք պատճառներով: Այսպես՝ դրա օգնությամբ կարելի է ուղղել միայն համացանցի միջոցով հրապարակվող կամ ուղարկվող տեքստերը (էլեկտրոնային նամակներ, հաղորդագրություններ, սոցիալական կայքերում արված հրապարակումներ): Դրանից բացի՝ ծրագիրը կիրառելի է միայն Firefox դիտարկիչի համար, մինչդեռ համացանցից օգտվողների ճնշող մեծամասնությունն օգտագործում է Google և Yandex որոնողական համակարգերը: Դա նշանակում է, որ նրանք պարզապես չեն կարող օգտվել տեքստի սրբագրման այս հավելվածից: Էլեկտրոնային սրբագրման նոր համակարգը, փոխարենը, հնարավոր կլինի օգտագործել բոլոր դիտարկիչներով և որոնողական համակարգերով:

Հայերեն տեքստերի սրբագրման համակարգի ստեղծման համար նախնառաջ պետք է հստակորեն որոշել դրանում ընդգրկվելիք բառապաշարի ծավալը՝ նպատակ ունենալով հայերենի բառապաշարի համաժամանակյա վիճակն արտացոլելը: Որքան մեծաքանակ լինի բառապաշարը, այնքան ամբողջական և լիակատար կլինի տեքստային սխալների ուղղումը: Կարծում ենք՝ նոր ստեղծվելիք համակարգի բառային հենքը պետք է կազմեն հայերենի առավել տարածված և ուղենշային համարվող բառարանները: Տեքստի սրբագրման մյուս ծրագրերի հեղինակները ևս կարևորում են այս հանգամանքը, սակայն իրականում նրանց ընդգրկած բառապաշարը փոքրաթիվ է: Օրինակ՝ Armenian PowerSpell-ի հեղինակների կարծիքով իրենց համակարգը բավական հարուստ է, քանի որ բաղկացած է 1,400,000 բառից: Հայերենի հարուստ բառազանձը չի հատում կես միլիոն բառի սահմանը, ուստի այս դեպքում, հասկանալի է, որ 1,400,000-ը ոչ թե բառերի, այլ բառաձևերի քանակն է: Եթե այդ թվից դուրս գան թեքվող խոսքի մասերի դեպքում հարցուցային շարքի մյուս անդամները, կստացվի, որ այս բառաքանակը կրճատվում է մի քանի անգամ:

Համակարգչային ծրագրերի կողմից բնական լեզվի արդյունավետ մշակումն ամենից առաջ պայմանավորված է ոչ միայն մեծածավալ բառապա-

շարի առկայությամբ, այլն լեզվի մակարդակային բոլոր միավորների առավել ամբողջական ու լիակատար ձևային նկարագրությամբ: Սա նշանակում է, որ լեզվի բառապաշարի, ինչպես նաև քերականության ձևայնացումը առաջին և կարևորագույն նախապայմանն է սրբագրման նոր համակարգի հաջողվածության հարցում: Ներկայացման ձևաչափի և ընդգրկելիք բառապաշարի տեսանկյունից շահեկանորեն տարբերվելով հայալեզու սրբագրման մյուս ծրագրերից՝ լեզվական սխալների ուղղման անհամեմատ բարձր մակարդակն է, որով էլեկտրոնային նոր համակարգը կտարբերվի այս և հետագա շրջանում ստեղծվելիք մյուս ծրագրերից:

Համակարգի առաջնային նպատակը լեզվի մակարդակային տարբեր միավորներում եղած լեզվագործածության տարատեսակ սխալների ուղղումն է:

ա) Ուղղագրական սխալների ուղղում: Ընդհանուր առմամբ՝ ցանկացած սրբագրման ծրագրի առաջնային խնդիրը ուղղագրական սխալների ուղղումն է: Գրավոր խոսքում ուղղագրական սխալները բազմաթիվ են, որոնք ընդհանրացնելով կարելի է բաժանել հետևյալ տեսակների. ա) վրիպակներ, բ) բառի ճիշտ գրության չիմացությամբ պայմանավորված սխալներ, գ) բառի ճիշտ գրության՝ բոլորի կողմից ընդունելի տարբերակի բացակայություն: Հայերենի ուղղագրական սխալների ուղղման հետ կապված խնդիրներն առավելապես պայմանավորված են վերջին հանգամանքով, քանի որ շատ դեպքերում բառերի միասին, անջատ կամ գծիկով գրության, հատուկ անունների դեպքում՝ բաղադրիչների մեծատառով կամ փոքրատառով գրության առնչությամբ լեզվաբանների շրջանում չկա բոլորի կողմից ընդունելի մեկ միասնական մոտեցում:

բ) Քերականական սխալների ուղղում: Սրբագրման համակարգի լիարժեքության մյուս նախապայմանը տեքստում եղած քերականական սխալների ուղղումն է, որը շատ դեպքերում բավական բարդ կլինի համակարգի համար: Այդ բարդություններն ամենից առաջ պայմանավորված են համակարգչային ծրագրերի՝ բնական լեզուն ոչ լիարժեք կերպով մշակելու կարողությամբ և հայերենի լիակատար ձևային նկարագրության բացակայությամբ, որը հիմնականում պայմանավորված է քերականական կարգերի առատությամբ և դրանց արտահայտման ձևերի բազմազանությամբ: Դրանից բացի՝ քերականական սխալ կառույցների բացահայտման ժամանակ լրացուցիչ դժվարություն կառաջացնի այն հանգամանքը, որ շատ դեպքերում բնական լեզվում քերականական միևնույն իմաստը կարող է արտահայտվել մեկից ավելի գույքահեռ ձևերով, իսկ ծրագրավորման լեզվով ոչ միշտ է հնարավոր դա տարբերակել:

զ) Կետադրական սխալների ուղղում: Տեքստի լեզվական որակը մեծ չափով պայմանավորված է ոչ միայն բառերի ճիշտ ուղղագրությամբ, քերականական համապատասխան կառույցների ընտրությամբ, այլև կետադրությամբ: Տեքստերում եղած կետադրական սխալները բազմաթիվ են, որոնց պատճառով որոշ դեպքերում նույնիսկ հնարավոր է ոչ ճիշտ ընկալել նախադասության կամ պարբերության իմաստը: Այս է պատճառը, որ սրբագրման համակարգում մեծապես կարևորվում ենք կետադրական սխալների ուղղման հնարավորությունը:

դ) Ոճական սխալների ուղղում: Էլեկտրոնային սրբագրման համակարգի հետագա զարգացման տեսանկյունից առավել հեռահար նպատակ ենք համարում մեքենական եղանակով տեքստի ոճական սխալների ուղղումը, առօրյա-խոսակցական, բարբառային բառերի ու արտահայտությունների փոխարինումը գրական տարբերակներով և այլն:

Էլեկտրոնային սրբագրման համակարգում էապես կարևորում ենք տեքստում լեզվական սխալների ուղղման եղանակը, որը կարելի է անել երկու միջոցով.

ա) համակարգն ավտոմատ կերպով կուղղի տեքստը, այնուհետև այլ ձևաչափով կամ գունավորմամբ կներկայացվեն ուղղված սխալները,

բ) տեքստի մուտքագրումից հետո համակարգը կընդգծի բոլոր սխալները՝ օգտվողին առաջարկելով ուղղման երեք-չորս ամենահավանական տարբերակները, որից հետո օգտվողը կընտրի համապատասխան ճիշտ ձևը:

Կարծում ենք՝ երկրորդ միջոցի կիրառությունն առավել նպատակահարմար է, քանի որ քիչ չեն լինի դեպքերը, երբ բառի գործածության համատեքստը օբյեկտիվորեն չհասկանալու կամ բառիմաստը չընկալելու պատճառով համակարգը սխալ կուղղի բառը: Այսպես՝ համանունները, որոնք բնական լեզվում ունեն տարբեր իմաստներ, բայց չունեն ձևային տարբերակվածություն, համակարգչային ծրագիրը որևէ կերպ միմյանցից չի կարող տարբերակել: Ավելին՝ երբեմն բառիմաստային տարբերություններով պայմանավորված՝ դրանք ձևաբանական փոփոխությունների են ենթարկվում տարբեր եղանակով: Օրինակ՝ **սեր** գոյականը զգացմունք իմաստն արտահայտելիս այլ ձևով է հոլովվում, կաթի **սերը**՝ այլ: Այս դեպքում համակարգը չի կարող ինքնուրույն ընտրել համապատասխան բառաձևը և ուղղել՝ համանուններն ըստ իմաստի տարբերակել չկարողանալու պատճառով:

Էլեկտրոնային սրբագրման համակարգի ստեղծման լեզվաբանական և ծրագրային կողմին առնչվող խնդիրները բազմազան են, դրանց լուծման եղանակները՝ ևս: Համեմատաբար ավարտունության ցանկացած փուլում էլ այն հսկայական նշանակություն կունենա հայախոս հանրության երկու հատ-

վածների համար հավասարապես: Այնուամենայնիվ, հարկ է ընդգծել համակարգի կարևորությունը ամենից առաջ արևելահայերենով գրված ամենատարբեր բնույթի տեքստերը կանոնարկելու, դրանցում եղած լեզվական տարբնույթ սխալներն ուղղելու համար: Պակաս կարևոր չի լինի դրա նշանակությունը նաև արևմտահայերի համար, որոնք հատկապես ուղղագրական համակարգերի տարբերություններով պայմանավորված, ոչ միշտ են կարողանում ինչպես ճիշտ ուղղագրությամբ արևելահայերեն տեքստեր գրել, այնպես էլ առհասարակ լիարժեք կերպով կիրառել հայերենի քերականական կանոնները: Կարելի է փաստել, որ սրբագրման համակարգից օգտվողների թիվը տեսականորեն որևէ սահմանափակում չի ունենա, քանի որ դա լիարժեք կերպով կարող է սպասարկել ողջ հայախոս հանրությանը, և ավելին՝ այն օտարերկրացիներին, որոնք բավարար լեզվական գիտելիքներ չունենալու պատճառով չեն կարող կառուցել իրենց գրավոր խոսքը պատշաճ մակարդակով:

Տեքստի ավտոմատ սրբագրման համակարգի ստեղծումը էապես կնպաստի ինչպես լեզվաբանական հետազոտությունների բնագավառում համակարգչային տեխնիկայի կիրառությանը, համակարգչային ծրագրերի միջոցով տեքստերում եղած լեզվական սխալների վերացմանը, այնպես էլ լայն առումով հայերենի՝ որպես համացանցային լեզվի հաջող գործառնությանը:

**Hakobyan Frida – The purpose and objectives of the Armenian Electronic Proofreading Program.**– The purpose of the article is to present the principles of creating the Armenian Electronic Proofreading program. The necessity of a new Proofreading program is underlined, because even though there are a few spell and grammar checkers, they are incomplete, as they do not correct all types of mistakes. For the creation of a new Proofreading program, the importance of including rich vocabulary and formal description of the language are highlighted. The newly-created program is supposed to correct spelling, grammar, punctuation, and stylistic errors. The Armenian Electronic Proofreading program will highly contribute to the use of computer technology in the field of linguistic research, elimination of errors in texts through computer programs, as well as to the successful functioning of the Armenian language as a web language.

**Акопян Фрида – Цель и задачи программы Армянской электронной корректуры.**– Цель статьи представить принципы создания Армянской программы электронного корректуры. Подчеркнута необходимость новой программы корректуры, потому что, несмотря на наличие нескольких контрольных и грамматических программ, они неполны, поскольку они не исправляют все типы ошибок. Для создания новой программы корректуры

подчеркивается важность включения богатого словарного запаса и формального описания языка. Предполагается, что новая программа исправит орфографию, грамматику, пунктуацию и стилистические ошибки. Программа Армянской электронной корректуры будет в значительной степени способствовать использованию компьютерных технологий в области лингвистических исследований, устранению языковых ошибок в текстах с помощью компьютерных программ, а также успешному функционированию армянского языка в качестве веб-языка.