

ՄԵՐԻ ՍԱՐԳՍՅԱՆ

**ՀԱՅԵՐԵՆ ԷԼԵԿՏՐՈՆԱՅԻՆ
ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳ**

armspell



**The National Academy of Sciences of the Republic of
Armenia**

H.Acharian Institute of Language

**Национальная академия наук Республики Армения
Институт языка имени Р. Ачаряна**

MERI SARGSYAN

**ARMENIAN ELECTRONIC PROOFREADING
SYSTEM**

МЕРИ САРГСЯН

**АРМЯНСКАЯ ЭЛЕКТРОННАЯ СИСТЕМА
КОРРЕКТУРЫ**

Yerevan – 2021 – Ереван

**ՀԱՅԱՍՏԱՆԻ ՀԱՆՐԱՊԵՏՈՒԹՅԱՆ
ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԶԳԱՅԻՆ ԱԿԱԴԵՄԻԱ
Հ. ԱՃԱՌՅԱՆԻ ԱՆՎԱՆ ԼԵԶՎԻ ԻՆՍՏԻՏՈՒՏ**

ՄԵՐԻ ՍԱՐԳՍՅԱՆ

**ՀԱՅԵՐԵՆ ԷԼԵԿՏՐՈՆԱՅԻՆ
ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳ**

ԵՐԵՎԱՆ



2021

ՀՏԴ 811.19:004
ԳՄԴ 81.51+32.97
Ս 259

Տպագրվում է ՀՀ ԳԱԱ Հ. Աճառյանի անվան լեզվի ինստիտուտի
գիտական խորհրդի որոշմամբ:

*Հեղագրությունն իրականացվել է ՀՀ ԿԳՆ գիտության
կոմիտեի տրամադրած ֆինանսավորմամբ՝ 19YR-6B070
ծածկագրով գիտական թեմայի շրջանակներում:*

Խմբագիր՝ Բ.Գ.Դ., պրոֆեսոր Վ. Գ. Համբարձումյան

Սարգսյան Մերի Արթուրի

Ս 259 Հայերեն էլեկտրոնային սրբագրման համակարգ/ Մ.
Ա. Սարգսյան.- Եր.: «Տիր» հրատ., 2021.- 100 էջ:

«Հայերեն էլեկտրոնային սրբագրման համակարգ» ուսումնասիրությունը նվիրված է էլեկտրոնային սրբագրման համակարգերի քննությանը: Ներկայացվում են հայալեզու և օտարալեզու սրբագրման համակարգերի առավելությունները և թերությունները, ինչպես նաև ուսումնասիրության արդյունքում ստեղծված սրբագրման նոր համակարգը, որը գործում է համացանցային առանձին կայքում՝ **armspell.am** վեբ հասցեի ներքո: Այս համակարգի ստեղծումը կարող է նպաստել 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: Այն մեծ հեռանկարներ է բացում ինչպես տեսական-լեզվաբանական, այնպես էլ գործնական-կիրառական խնդիրների լուծման համար: Ծրագիրը կարող է օգտակար լինել թե՛ Հայաստանում, թե՛ Հայաստանից դուրս գտնվող ցանկացած մարդու համար՝ անկախ սեռից, տարիքից, կրթական մակարդակից, բնակության վայրից:

ՀՏԴ 811.19:004
ԳՄԴ 81.51+32.97

ISBN 378-9939-865-91-1

© Սարգսյան Մերի Արթուրի, 2021

ԲՈՎԱՆԴԱԿՈՒԹՅՈՒՆ

ԱՌԱՋԱԲԱՆ	6
ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԵՐԸ՝ ՈՐՊԵՍ ՏԵԽՆԻԿԱԿԱՆ ՆՈՐ ԼՈՒԾՈՒՄՆԵՐ	11
ՕՏԱՐԱԼԵԶՈՒ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԵՐ	22
ՀԱՅԱԼԵԶՈՒ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԵՐ	33
ARMSPELL ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԻ ԿԱԶՄՈՒԹՅԱՆ ՍԿԶԲՈՒՆՔՆԵՐԸ	48
ARMSPELL ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԻ ԱՌԱՆՁՆԱՀԱՏԿՈՒԹՅՈՒՆՆԵՐԸ	54
ARMSPELL ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԻ ՇՏԵՄԱՐԱՆԻ ԱՌԱՆՁՆԱՀԱՏԿՈՒԹՅՈՒՆՆԵՐԸ	59
ARMSPELL ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԻ ԿԻՐԱՌՈՒԹՅԱՆ ՈՒՂԵՑՈՒՅՑ	68
SUMMARY	82
ՕԳՏԱԳՈՐԾՎԱԾ ԳՐԱԿԱՆՈՒԹՅԱՆ ՑԱՆԿ	91

ԱՌԱՋԱԲԱՆ

Այսօր համացանցը ողողված է լեզվական տարատեսակ սխալներով, գռեհկաբանությունները, ժարգոնային արտահայտությունները, փողոցային խոսքը մուտք են գործել պետհաստատություններ, եթեր, մամուլ, մեր առօրյա: Օրեցօր մեր հասարակության մի ստվար հատված սկսում է դա դիտարկել որպես նորմ: Բացի այդ՝ համատարած լատինատառ գրելու միտումը մեկ այլ հետընթաց է առաջ բերում. մարդիկ դժվարանում են հայատառ գրագետ գրել: Եվ բնականաբար այս համատեքստում օրհասական է հայալեզու խմբագրիչների առկայությունը: Ինչու՞ խմբագրիչների: Ինչպես գիտենք, տեխնոլոգիաների զարգացման մեջ մեծ խթան է լինում մրցակցությունը, այս ոլորտը ևս բացառություն չէ, այստեղ ևս անհրաժեշտ է առողջ մրցակցություն: Որքան շատ լինեն խմբագրման և սրբագրման ծրագրերը, համակարգերը, այնքան ավելի որակյալ և արդյունավետ կդառնան մրցակցելու շնորհիվ:

Էլեկտրոնային սրբագրման համակարգերը ծրագրեր են, որոնց միջոցով կարելի է ուղղել տեքստի

լեզվական սխալները: Օտարալեզու սրբագրման համակարգեր ստեղծվել են դեռևս անցյալ դարում: Սկզբնական շրջանում դրանք ուղղում էին միայն ուղղագրական սխալներ, այնուհետև ծրագրային հնարավորությունները մեծացան, և սկսեցին ուղղել նաև քերականական սխալները: Այսօր օտարալեզու սրբագրման համակարգերը զարգացման բավական բարձր մակարդակ ունեն. դրանք կարողանում են ուղղել ոչ միայն ուղղագրական և քերականական սխալները, այլև՝ ոճական և կետադրական: Ավելին՝ ոլորտում առկա են անգամ բազմալեզվյան սրբագրման համակարգեր, որոնք հնարավորություն են տալիս ուղղելու մի քանի լեզուներով տեքստերի սխալները:

«Հայերեն էլեկտրոնային սրբագրման համակարգ» ուսումնասիրությունը նվիրված է էլեկտրոնային սրբագրման համակարգերի քննությանը: Ներկայացվում են հայալեզու և օտարալեզու սրբագրման համակարգերի առավելությունները և թերությունները, ինչպես նաև ուսումնասիրության արդյունքում ստեղծված սրբագրման նոր համակարգը, որը գործում է համացանցային առանձին կայքում:

ArmSpell հայերեն էլեկտրոնային սրբագրման համակարգը ստեղծվել է ՀՀ ԿԳՄՍ ԳԿ «Երիտասարդ գիտաշխատողների հետազոտությունների աջակցության» ծրագրի՝ «19YR-6B070» ծածկագրով «Հայերեն էլեկտրոնային սրբագրման համակարգ» գիտահետազոտական թեմայի շրջանակում: Ծրագրի ղեկավարը ՀՀ ԳԱԱ Հ. Աճառյանի անվան լեզվի ինստիտուտի ավագ գիտաշխատող, բանասիրական գիտությունների թեկնածու, դոցենտ Մերի Սարգսյանն է: Թեմայի շրջանակում հրատարակվել են գիտական հոդվածներ՝ տեղի և արտասահմանյան ամսագրերում¹:

¹ *Տե՛ս* M. A. Sargsyan, Some Observations on Armenian Electronic Proofreading Systems, "World Science", №6 (34), Vol. 8, June 2018, Warsaw, Poland, pp. 21-24, DOI:https://doi.org/10.31435/rsglobal_ws/12062018/5888; Մ. Ա. Սարգսյան, Հայերեն էլեկտրոնային սրբագրման առկա համակարգերի քննություն, Բանբեր Երևանի Վ. Բրյուսովի անվան պետական լեզվահասարակագիտական համալսարանի, 1(44), Երևան, 2018, էջ 598-604; Մ. Ա. Սարգսյան, Հայալեզու տեքստերի էլեկտրոնային նոր սրբագրիչի սկզբունքները, «Լեզու և լեզվաբանություն», հ. 2 (23), 2020, Երևան, էջ 10-17; Մ. Ա. Սարգսյան, էլեկտրոնային սրբագրիչի տվյալների շտեմարանի բառապաշարի ձևային նկարագրության խնդիրների մասին, «Բանբեր Երևանի Վ. Բրյուսովի անվան պետական լեզվահասարակագիտական համալսարանի», 2021, հ. 1 (58), էջ 229-237:

«Հայերեն էլեկտրոնային սրբագրման համակարգ» գիտահետազոտական թեմայի շրջանակներում ստեղծվել է սրբագրման մի համակարգ, որի միջոցով փորձ է արվում ամբողջական և լիարժեք կերպով սրբագրել արդի արևելահայերենով գրված ցանկացած տեքստ: Հանրության լայն շրջանակներին հասանելի դարձնելու նպատակով սրբագրման համակարգը գործում է համացանցային առանձին կայքում՝ <https://armspell.am/hy> վեբ կայքի հասցեի ներքո, որը գրանցված է որոնողական բոլոր համակարգերում²: Կիրառման տեսակետից այն պարզ է և արագագործ, քանի որ դրանից օգտվելու համար որևէ ծրագիր ներբեռնելու անհրաժեշտություն չկա:

Այս համակարգի ստեղծումը կարող է նպաստել ինչպես լեզվաբանական հետազոտությունների բնագավառում համակարգչային տեխնիկայի կիրառությանը, համակարգչային ծրագրերի միջոցով տեքստերում եղած լեզվական սխալների վերացմանը, այնպես էլ լայն առումով հայերենի՝ որպես համացանցային լեզվի հաջող գործառությանը: Այն մեծ հեռանկարներ է բացում ինչպես տեսական-

² «Հայերեն էլեկտրոնային սրբագրման համակարգ», <https://armspell.am/hy>:

լեզվաբանական, այնպես էլ գործնական-կիրառական
խնդիրների լուծման համար: Ծրագիրը կարող է
օգտակար լինել թե՛ Հայաստանում, թե՛ Հայաստանից
դուրս գտնվող ցանկացած մարդու համար՝ անկախ
սեռից, տարիքից, կրթական մակարդակից, բնա-
կության վայրից:

ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԵՐԸ՝ ՈՐՊԵՍ ՏԵԽՆԻԿԱԿԱՆ ՆՈՐ ԼՈՒԾՈՒՄՆԵՐ

Տեքստը սրբագրելու ավտոմատ համակարգերի ստեղծումը բնական լեզվի մշակման տեխնիկայի ամենահին կիրառություններից է, քանի որ, ինչպես նշում է Ռ. Միթոնը, ավտոմատ սրբագրման առաջին համակարգերը ի հայտ են եկել 50-ականների վերջին³: Ցանկացած սրբագրիչ կատարում է երկու կարևոր գործառույթ՝ հաջորդաբար՝ սկզբում հայտնաբերում է, ապա ուղղում ուղղագրական սխալները: Հայտնաբերելու և ուղղելու մեթոդներն աշխատում են երեք եղանակով.

1. ***սխալի հայտնաբերում***, երբ համակարգը հայտնաբերում է ուղղագրական սխալ ունեցող առանձին բառերը կամ օտար բառերը,

2. ***նախապես որևէ կոնտեքստում հանդիպած սխալի հայտնաբերում***, երբ համակարգը հայտնաբերում է նախկինում որևէ կոնտեքստում եղած բա-

³ St'u Mitton, R.: Ordering the suggestions of a spellchecker without using context, Natural Language Engineering, (2009):

ոի սխալը՝ անկախ նրանից, թե այս բառը ինչ բառերով է շրջապատված,

3. համատեքստային սխալի հայտնաբերում և ուղղում, երբ համակարգը հայտնաբերում և ուղղում է սխալները՝ հաշվի առնելով համատեքստը. այս դեպքում ուղղվում են նույնիսկ այն բառերը, որոնք կան լեզվում, սակայն տվյալ համատեքստում անտեղի է դրանց կիրառությունը⁴:

Ծրագրային ապահովման մեջ սրբագրիչը կարելի է բնորոշել որպես ծրագրային ապահովման գործառույթ, որը ստուգում է տեքստի տառասխալները: Ուղղագրության ստուգման գործառույթները հաճախ ներդրված են ծրագրային ապահովման կամ ծառայությունների մեջ, ինչպիսիք են, օրինակ, էլեկտրոնային բառարանը կամ որոնողական համակարգը:

Սովորաբար սրբագրիչներն իրականացնում են հետևյալ հիմնական գործընթացները.

- Այն սկանավորում է տեքստը և առանձնացնում դրանում պարունակվող բառերը:

⁴ Տե՛ս Kukich, K.: Techniques for Automatically Correcting Words in Text. R. Mitton, Ordering the suggestions of a spellchecker without using context, (2009).

- Այնուհետև յուրաքանչյուր բառը համեմատում է ճիշտ գրված բառերի ցանկի հետ (այսինքն՝ բառարանի): Դա կարող է լինել ընդամենը բառերի ցանկ, կամ կարող է լինել նաև լրացուցիչ տեղեկություններ պարունակող շտեմարան, օրինակ՝ կարող է ընդգրկել գծագրեր կամ բառապաշարային և քերականական հատկանիշների տվյալներ ևն:

- Հաջորդ քայլով կազմվում է արդեն ալգորիթմ՝ հիմնված լեզվական տվյալների վրա: Սա ձևաբանական փոփոխությունները մշակելու և ստուգելու համար է: Նույնիսկ ձևաբանորեն թեթև փոփոխություններ կրող լեզուների դեպքում անպայման պետք է հաշվի առնել այս քայլը: Օրինակ՝ անգլերեն ուղղագրիչը պետք է հաշվի առնի միևնույն բառի տարբեր ձևեր, ինչպիսիք են հոգնակիի ձևերը, բառային զուգաձևությունները ևն: Կցական լեզուների դեպքում այս գործընթացը շատ ավելի բարդ է լինում և աշխատատար:

Սովորաբար սրբագրման համակարգերն այնպես են մշակվում, որ ծրագրից օգտվողները կարողանում են հաստատել կամ մերժել առաջարկվող տարբերակը: Սրբագրիչներ մշակելիս սխալ գրված բառերի ճիշտ ուղղագրությունը գտնելու համար հաճախ կի-

րառում են այնպիսի ալգորիթմներ, ինչպիսին է **Լևենշտեյնի հեռավորությունը**⁵: Որոշ սրբագրիչներում կիրառվում է այլընտրանքային տարբերակ՝ վիճակագրական տեղեկատվություն, օրինակ՝ **n-grams**, որը ճիշտ գրված բառերի փոխարեն ճանաչում է սխալ գրված բառերը: Պետք է նշել, որ այս տարբերակը կիրառելու համար անհրաժեշտ է լինում վիճակագրական տեղեկատվություն հավաքել, որը երբեմն որպես թերություն է դիտարկվում: Չնայած ակնհայտ առավելություններ էլ կան, ինչպես, օրինակ՝ գործարկման համար ավելի քիչ ժամանակ է պահանջվում, բացի այդ՝ կարելի է ուղղել նաև շտեմարանում չընդգրկված բառերի ուղղագրական սխալները⁶:

Որոշ դեպքերում էլ էլեկտրոնային սրբագրիչների հիմքում դրված է լինում **ուղղագրական սխալների կայուն ցանկ**, և ուղղելու առաջարկները ձևավորվում են այդ սխալների համար միայն: Սա, դժվար չէ տես-

⁵ Perner, Petra (2010-07-05). Advances in Data Mining: Applications and Theoretical Aspects: 10th Industrial Conference, ICDM 2010, Berlin, Germany, July 12-14, 2010. Proceedings. Springer Science & Business Media. ISBN 978-3-642-14399-1.

⁶ U.S. Patent 6618697, Method for rule-based correction of spelling and grammar errors.

նել, որ բավական սահմանափակ և ոչ ճկուն տարբերակ է:

Առաջին սրբագրիչն ի հայտ է եկել 1961 թվականին, հեղինակը ամերիկացի ծրագրավորող **Լես Էռնեստը** է: Նրա ստեղծած սրբագրիչը պարունակում էր 10 000 բառ ունեցող ցանկ⁷: Հետագայում՝ 1971 թ., Էռնեստի ուսանող **Ռալֆ Գորինը** Սթենֆորդի համալսարանի Արհեստական բանականության լաբորատորիայում ստեղծեց ուղղագրության ստուգման առաջին ծրագիրը, որը գրված էր որպես կիրառական ծրագիր (ոչ թե հետազոտական)՝ անգլերեն տեքստերը ստուգելու համար⁸: Գորինի ստեղծած սրբագրիչը՝ **SPELL**-ը, բառերի ցանկից փնտրում և գտնում էր մեկ տառով կամ մի քանի տառերով տարբերվող բառերը և ներկայացնում օգտվողին: Ինչպես Սթենֆորդի արհեստական բանականության լաբորատորիայի ծրագրերի մեծ մասը, այնպես էլ այս սրբագրիչը բաց մատչե-

⁷ Earnest, Les. "The First Three Spelling Checkers" (PDF). Stanford University. Archived from the original (PDF) on 22 October 2012. Retrieved 10 October 2011.

⁸ Peterson, James (December 1980). *Computer Programs for Detecting and Correcting Spelling Errors* (PDF). Retrieved 2021-02-18.

լիությամբ ներկայացվեց հանրությանը՝ դրանով իսկ մեծ տարածում ու հեղինակություն գտնելով⁹:

Հետագայում SPELL-ը, նրա ալգորիթմները և տվյալների շտեմարանի կառուցվածքը հիմք դարձան մեկ այլ սրբագրիչի՝ **Unix ispell** ծրագրի համար: Զորջթաունի համալսարանի վեց լեզվաբաններից բաղկացած խումբը մշակեց IBM կորպորացիայի համար ուղղագրության ստուգման առաջին համակարգը¹⁰: 1981 թվականին **Հենրի Կուչերան** ստեղծեց սրբագրիչ Digital Equipment Corp-ի VAX մեքենաների համար¹¹:

Այսպիսով՝ 1970-ականների վերջին էլեկտրոնային առաջին սրբագրիչներն արդեն լայնորեն հասանելի էին:

Անձնական համակարգիչների համար առաջին սրբագրիչներն ի հայտ եկան 1980 թվականին: Այդպիսի մի սրբագրիչ էր «**WordCheck**»-ը Commodore

⁹ Earnest, Les. *Visible Legacies for Y3K* (PDF). Archived from the original (PDF) on 2011-07-20. Retrieved 2011-02-18.

¹⁰ "Georgetown U Faculty & Staff: The Center for Language, Education & Development". Archived from the original on 2009-02-05. Retrieved 2008-12-18.

¹¹ Harvey, Charlotte Bruce (May–June 2010). "Teaching Computers to Spell (obituary for Henry Kučera)". *Brown Alumni Magazine*. p. 79.

համակարգերի համար, որը թողարկվեց 1980 թվականի վերջին՝ գովազդի տպագրության ժամանակ¹²: Նախավինդոուզյան համակարգիչների համար նախատեսված այս սրբագրիչները ինքնուրույն ծրագրեր էին, որոնցից շատերը կարող էին գործարկվել TSR ռեժիմով:

Այնուամենայնիվ, այդ ծրագրերը կարճատև կյանք ունեցան, քանի որ 1980-ականների կեսերին **WordStar**-ի և **WordPerfect**-ի նման բառամշակման փաթեթների հայտնի մշակողները իրենց փաթեթներում ներառել էին սրբագրիչներ, որոնք հիմնականում արտոնագրված էին, իսկ դա էլ նպաստեց, որպեսզի այդ ծրագրերն արագ տարածում գտնեն ոչ միայն Եվրոպայում, այլև ասիական շատ երկրներում:

Չնայած առկա դժվարություններին, որոնք հիմնականում պայմանավորված էին կցական լեզուների ձևաբանական պատկերը ներկայացնելու խնդրով, այս ընկերությունները որդեգրել էին մարքեթինգային ռազմավարության գլոբալ մոտեցումը և ջանում էին

¹² Advertisement (January 1981). "Micro Computer Industries, Ltd" (PDF). *Compute! Magazine, Issue 8, Vol. 3, No. 1*. p. 119.

իրենց ծրագրերը տեղայնացնել հնարավորինս շատ ազգային շուկաներում:

Երբ **Apple**-ը Mac OS X-ի համար մշակեց **«ամբողջ համակարգի սրբագրիչ»**, դա եզակի էր իր տեսակի մեջ: Ընկերությունը կարծում էր, որ անհրաժեշտություն չկա յուրաքանչյուր առանձին ծրագրի համար առանձին սրբագրիչ ստեղծել: Հետևաբար ստեղծեց մեկը, որով օպերացիոն համակարգը կատարում էր ամբողջ համակարգի ուղղագրական սխալների ստուգում¹³:

Առաջին սրբագրիչները եղել են **«ստուգիչներ»**՝ **«ուղղիչներ»** փոխարեն: Նրանք ոչ մի առաջարկ չէին ներկայացնում սխալ գրված բառերի փոխարեն: Այն օգտակար էր տառասխալների համար, բայց այնքան էլ օգտակար չէր տրամաբանական կամ հնչյունական սխալների դեպքում:

Մշակողների առջև ծառացած ամենամեծ խնդիրը սխալ գրված բառերի փոխարեն օգտակար առաջարկներ ներկայացնելու դժվարությունն էր: Այն

¹³ David Pogue (2009). Mac OS X Snow Leopard: The Missing Manual; David Pogue (2015). Switching to the Mac: The Missing Manual.

պահանջում էր ձևայնացնել բառերը և կիրառել օրինաչափությունների համապատասխան ալգորիթմներ:

Հասկանալի է, որ ուղղագրության ստուգման ծրագրի հիմքում ընկած բառարանները որքան մեծ և ընդգրկուն են, այնքան լավ է աշխատում ծրագիրը, քանի որ այդ դեպքում ճիշտ բառերը չեն դիտարկվում որպես սխալ: Գործնականում, սակայն, անգլերենի համար սրբագրիչի շտեմարանի արդյունավետ չափը համարվում է 90,000 միավորը: Եթե շտեմարանի միավորների թիվը գերազանցում է այդ շեմը, մեկ այլ վտանգ է առաջանում. սխալ գրված բառերը շփոթվում են այլ բառերի հետ և չեն դիտարկվում որպես սխալ:

Առաջին **MS-DOS սրբագրիչները** հիմնականում ընդգրկված էին բառի մշակման փաթեթներում: Փաստաթուղթը պատրաստելուց հետո օգտատերը սկանավորում էր տեքստը՝ փնտրելով ուղղագրական սխալներ:

Հետագայում, սակայն, առաջարկվեց այնպիսի փաթեթ, ինչպիսին է **Oracle-ի CoAuthor-ը**, որը թույլ տվեց օգտվողին դիտել արդյունքները փաստաթուղթը

մշակելուց հետո և ուղղել միայն այն բառերը, որոնք հայտնի էին որպես սխալ: Երբ հիշողության և ծրագրային հզորությունը մեծացավ, ուղղագրության ստուգումն իրականացվում էր արդեն հետին պլանում՝ ինտերակտիվ եղանակով, ինչպես դա արվում էր 1987 թվականին թողարկված **Sector Software-ի Spellbound** ծրագրի և **Microsoft Word-ի՝ Word 95-ից** սկսած ծրագրերի դեպքում:

Վերջին տարիներին սրբագրիչներն աստիճանաբար ավելի բարդ են դարձել. դրանց մի մասն այժմ ունակ է ճանաչելու քերականական պարզ սխալները: Այնուամենայնիվ, նույնիսկ լավագույն դեպքում դրանք հազվադեպ են ընկալում տեքստի բոլոր սխալները (օրինակ՝ և՛ համանուն բառերը, և՛ նորաբանությունները, և՛ օտար բառերը նշում են որպես ուղղագրական սխալ ունեցող բառեր ևն):

Չնայած այդ ամենին՝ սրբագրիչները կիրառական մեծ նշանակություն ունեն, համարվում են օտար լեզվով ճիշտ գրելու օգտակար ծրագրեր, որոնց շնոր-

հիվ օտարախոսները հեշտությամբ կարողանում են տեսնել իրենց սխալները և շտկել¹⁴:

¹⁴ Banks, T. (2008). *Foreign Language Learning Difficulties and Teaching Strategies*. (pp. 29). Master's Thesis, Dominican University of California. Retrieved 19 March 2012.

ՕՏԱՐԱԼԵԶՈՒ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԵՐ

Էլեկտրոնային սրբագրման համակարգերը ծրագրեր են, որոնց միջոցով կարելի է ուղղել տեքստի լեզվական սխալները: «Հայերեն էլեկտրոնային սրբագրման համակարգ» գիտահետազոտական թեման իրականացնելիս ուսումնասիրել ենք մի շարք սրբագրման համակարգեր՝ օտարալեզու և հայալեզու:

Օտարալեզու սրբագրման համակարգեր ստեղծվել են դեռևս անցյալ դարում: Սկզբնական շրջանում դրանք ուղղում էին միայն ուղղագրական սխալներ, այնուհետև ծրագրային հնարավորությունները մեծացան, և սկսեցին ուղղել նաև քերականական սխալները:

Օտարալեզու առաջին սրբագրման համակարգերից է **Writer's Workbench**-ը: Այն ծրագրային փաթեթ է, որը մշակվել է Unix օպերացիոն համակարգի համար *Լորինդա Չերիի* և *Նինա Մակդոնալդի* կողմից¹⁵:

¹⁵ Smith, Charles R.; Kathleen E. Kiefer; Patricia S. Gingrich (1 July 1984). "Computers come of age in writing instruction". Computers

Դա, թերևս, քերականության ամենավաղ սրբագրիչն էր, որը լայն կիրառություն ստացավ **Unix** համակարգերում¹⁶: Writer's Workbench համակարգի հնարավորություններից կարելի է հիշատակել հետևյալները. **diction** գործիքը հնարավորություն է տալիս ուղղելու սխալ գրված բառերն ու արտահայտությունները, **style** գործիքը հնարավորություն է տալիս վերլուծել տեքստը, վիճակագրական տվյալներ ստանալ նախադասության կառուցվածքի մասին, հայտնաբերել ոճական սխալները¹⁷:

Սրբագրման առաջին ծրագրերից հիշատակության արժանի է նաև **Grammatik**-ը, որը ստեղծվել է 1981 թվականին, այն ևս հաջողությամբ հայտնաբերում և ուղղում էր քերականական սխալները¹⁸: Հետա-

and the Humanities. Springer Netherlands. 18 (3): 215–224. doi:10.1007/BF0226 7225. S2CID 28762117.

¹⁶ Dale, Robert; Moisl, Hermann; Somers, Harold (2000). Handbook of Natural Language Processing. CRC Press. ISBN 9780824790004.

¹⁷ Smith, Charles R.; Kathleen E. Kiefer; Patricia S. Gingrich (1 July 1984). "Computers come of age in writing instruction". Computers and the Humanities. Springer Netherlands. 18 (3): 215–224. doi:10.1007/BF0 2267225. S2CID 28762117.

¹⁸ "Grammatik II". PC: The Independent Guide to IBM Personal Computers. Software Communications. 5: 190–199. 1986.

գայում սրբագրիչների հնարավորությունները ընդլայնվեցին, դրանք արդեն ուղղում էին ուղղագրական, քերականական, կետադրական սխալները: Այդպիսի ծրագրերից նշենք հետևյալները՝ ***Punctuation & Style, Correct Grammar, RightWriter***¹⁹:

Այսօր օտարալեզու սրբագրման համակարգերը զարգացման բավական բարձր մակարդակ ունեն. դրանք կարողանում են ուղղել ոչ միայն ուղղագրական և քերականական սխալները, այլև՝ ոճական և կետադրական²⁰: Ավելին՝ ոլորտում առկա են անգամ բազմալեզվյան սրբագրման համակարգեր, որոնք հնարավորություն են տալիս ուղղելու մի քանի լեզուներով տեքստերի սխալները: Հայտնիներից են հետևյալները՝ ***Grammarly, Ginger, Spell Check PlusPro, CorrectEnglish Complete*** և այլն: Այս ծրագրերն աշխատում են առցանց, որը այժմ բավական կիրառելի է և մատչելի է դարձնում ծրագրային թարմացումներից օգտվելը:

¹⁹ InfoWorld Media Group (28 October 1991). InfoWorld. InfoWorld Media Group, Inc. p. 68.

²⁰ Peterson James, . Computer Programs for Detecting and Correcting Spelling Errors (PDF). (Dec 1980), Retrieved 2011-02-18.

Կան ծրագրեր, որոնք կարելի է ավելացնել դիտարկչի մեջ՝ որպես հավելված (Chrome, Safari, Firefox, Edge): Կան նաև այնպիսի ծրագրեր, որոնք հնարավոր է ներբեռնել համակարգչի մեջ: Սովորաբար այդ ծրագրերը կիրառվում են որոշակի կոնկրետ օպերացիոն համակարգերի դեպքում, որը կարող ենք համարել ծրագրային թերություն, բացի այդ՝ այս դեպքում ծրագրային յուրաքանչյուր թարմացումից հետո պետք է ձեռք բերել ծրագրի նորացված տարբերակը, որը շատ հաճախ վճարովի է լինում կամ անհրաժեշտ են լինում լրացուցիչ թղթաբանական միջամտություններ, որը ևս դժվարացնում է դրանից օգտվելու գործընթացը:

Այս ոլորտում թերևս առաջատարն է **Grammarly** սրբագրման համակարգը: Այն իր հնարավորություններով ավելի ընդգրկուն է և արդյունավետ: Բացի այն, որ այս սրբագրիչը ուղղում է վերը նշված բոլոր տիպի սխալները, այն կարելի է կիրառել երկու տարբերակով՝ բեռնել համակարգիչ, ավելացնել **MS Word, Excell, PowerPoint** ծրագրային հավելվածներին և կիրառել այդտեղ, կամ օգտագործել առցանց՝ առանձին կայքում: Այս սրբագրիչն ունի նաև *բջջային հավելվածի* տարբերակ, որը ևս մեծ առավելություն է:

Ինչպես գիտենք, այժմ բջջային հավելվածները լիովին կամ լիարժեքորեն փոխարինում են համակարգչային ծրագրերին: Բնականաբար՝ համակարգչային ցանկացած ծրագրի բջջային հավելվածի առկայությունը կարող է դիտարկվել որպես ծրագրային լուրջ առավելություն և մրցունակության բարձր ցուցիչ: Օրինակ՝ **Grammarly** սրբագրիչի բջջային հավելվածը կիրառելի է և՛ iOS, և՛ Android օպերացիոն համակարգերի համար: Մեր ուսումնասիրությունները սկսելիս, իհարկե, էտալոնային համարել ենք հենց այս սրբագրիչը՝ իր հնարավորություններով, որը տարիների, մասնագիտական և ֆինանսական մեծ ռեսուրսների առկայության արդյունք է:

Ըստ տարբեր լեզուներով տեքստեր ուղղելու հնարավորության՝ սրբագրման համակարգերը լիում են միալեզվյան և բազմալեզվյան: **Միալեզվյան համակարգերը** հնարավորություն են տալիս ուղղել սխալները որևէ կոնկրետ լեզվի շրջանակում, **բազմալեզվյան ծրագրերի** հնարավորություններն ավելի ընդգրկուն են. դրանք ուղղում են մեկից ավելի լեզուներով գրված տեքստեր: Ի պատիվ ոլորտի մասնագետների լուրջ և պրպտուն աշխատանքի՝ այժմ ասպարեզում կան բազմալեզվյան բավական հաջող

սրբագրիչներ, ինչպես, օրինակ՝ **Grammarly, Ginger, Spell Check PlusPro, CorrectEnglish Complete, WhiteSmoke** և այլն:

Ինչպե՞ս են աշխատում սրբագրիչները կամ ինչպե՞ս են ուղղում լեզվական սխալները

Ցանկացած սրբագրման համակարգ կամ ծրագիր լեզվական սխալներն ուղղում է հետևյալ կերպ.

- տեքստը մեքենական եղանակով վերլուծվում է, յուրաքանչյուր միավոր դիտարկվում է որպես բառաձև,

- յուրաքանչյուր բառաձև համեմատվում է տվյալ սրբագրիչի հիմքում ընկած բառարանի հետ, հետևաբար՝ սրբագրիչը կարող է հաջող համարվել, եթե հիմքում ընկած բառարանում ներառված են բառի ոչ միայն ուղիղ, բառարանային ձևերը, այլև թեքված բոլոր ձևերը՝ բառաձևերը,

- երբեմն սրբագրիչի հիմքում ընկած է լինում հնարավոր լեզվական սխալների բառացանկ, և

սրբագրիչը սխալների ուղղումը կատարում է այդ ցանկի շրջանակում,

- համակարգը, հայտնաբերելով հնչյունական կազմի նմանություն, առաջարկում է սխալն ուղղելու ճիշտ տարբերակ կամ տարբերակներ,

- օգտվողը, տեսնելով համակարգի առաջարկած տարբերակը կամ տարբերակները, ընտրում է տարբերակներից որևէ մեկը և ուղղում սխալը,

- կան սրբագրիչներ, որտեղ օգտվողը կարող է համալրել սրբագրման համակարգի հիմքում ընկած շտեմարանը՝ ավելացնելով սխալն ուղղելու նոր տարբերակ, որը ներառված չէ շտեմարանում, սակայն սա շատ խնդրահարույց է, կարող է կամայական սխալների պատճառ դառնալ,

- կան սրբագրիչներ, որոնց հիմքում այնպիսի ալգորիթմներ են, որոնք լեզվական սխալները հայտնաբերում են՝ հաշվի առնելով բառի գործածության համատեքստը:

Էլեկտրոնային սրբագրման համակարգերի հիմքում բավական լուրջ ծրագրաշար է, որը ներառում է այնպիսի ալգորիթմներ, որոնք հնարավորություն են տալիս մշակել լեզվական նյութը, հայտնաբերել տեքս-

տում եղած լեզվական սխալները, առաջարկել ուղղելու ճիշտ տարբերակներ:

Ինչ վերաբերում է էլեկտրոնային սրբագրիչների լեզվաբանական կողմին, ապա դրա հաջող գործառման հիմքում ընկած է լեզվի լիարժեք ձևային նկարագրությունը: Որքան լիարժեք է ձևային նկարագրությունը, այնքան հաջող է սրբագրիչը: Բայցևայնպես, լեզվում կան այլևայլ կողմեր, որոնք ենթակա չեն ձևայնացման, հետևաբար՝ բոլոր սրբագրիչներն էլ՝ թե՛ հայալեզու, թե՛ օտարալեզու, թե՛ միալեզվյան, թե՛ բազմալեզու, ունեն նույն խնդիրը. կան լեզվական սխալներ, որոնք դեռևս հնարավոր չէ ուղղել գոյություն ունեցող որևէ սրբագրիչի օգնությամբ: Սակայն, պետք է նշել, որ դա չի վերաբերում **ուղղագրական սխալները ուղղելուն**:

Գրեթե բոլոր սրբագրիչները լիարժեք կատարում են այս գործառույթը: Կան որոշ բացառություններ, որոնք կապված են շտեմարանի սակավության հետ: Անշուշտ, լեզուն կենդանի օրգանիզմ է, շարունակ զարգանում է, հատկապես ակտիվորեն փոփոխվում ու համալրվում է բառապաշարը, ի հայտ են գալիս նոր բառեր, որևէ լեզու են մուտք գործում օտար բառեր,

որոնք հնարավոր չի լինում ներառել տվյալների շտեմարանում: Այսինքն՝ ուղղագրական որոշ սխալներ չհայտնաբերելու և չուղղելու առավել տարածված խնդիրը սա է: Երբեմն խնդիրը կապված է լինում համատեքստում բառի ոչ ճշգրիտ կիրառման հետ: Դեռևս ոչ բոլոր սրբագրիչներն են կարողանում լիարժեք վերլուծել համատեքստը և հայտնաբերել բառի ճիշտ կամ սխալ կիրառությունը՝ ըստ այդ համատեքստի:

Ինչ վերաբերում է **քերականական սխալները ուղղելուն**, ապա պետք է նշել, որ ասպարեզում չենք հանդիպել օտարալեզու որևէ սրբագրիչ, որը լիարժեք ուղղի քերականական բոլոր տիպի սխալները: Այսպես, օրինակ՝ սրբագրիչը կարող է ուղղել քերականական թվին առնչվող սխալները, բայց չի կարողանում ուղղել ածականի համեմատության աստիճանի սխալը:

Օտարալեզու սրբագրիչների գրեթե հիմնական մասը հաջողությամբ ուղղում է նաև տեքստի **կեղտահարական սխալները**, որոշ սրբագրիչներ, ինչպիսին **Grammarly**-ն է, հաջողությամբ ուղղում է նաև տեքստի **ոճական սխալները**: Իհարկե պետք է նշել, որ ծրագրային այս հնարավորությունը բաց մատչելիութ-

յամբ չէ, այլ վճարովի տարբերակում է հասանելի, այնուամենայնիվ, սա մեծ առավելություն է և մեծ օգնություն մասնագիտական տեքստի խմբագիրներին, որը տարբեր ռեսուրսների խնայողության հնարավորություն է տալիս: Այսպես, օրինակ՝ ոճական առումով խճողված է համարվում բարդ նախադասություններ պարունակող տեքստը, կամ համակարգը որոշ բառերի՝ ոճական առումով ավելի համապատասխան տարբերակներ է առաջարկում և այլն:

Եթե համեմատենք հայալեզու և օտարալեզու սրբագրման լավագույն համակարգերը, ապա կարող ենք արձանագրել հետևյալ պատկերը: Ինչպես նշեցինք, օտարալեզու սրբագրիչներից առաջատարը և լավագույնը Grammarly-ն է, հայալեզու գործող սրբագրիչներից համեմատաբար հաջող գործարկվում են Armenian PowerSpell և HySpell ծրագրերը: Բայց այնուամենայնիվ դեռևս հեռու են այս սրբագրիչները համեմատելու եզրերը: Ցավոք, հայալեզու սրբագրիչները դեռևս ուղղագրական և քերականական խնդիրներ լուծելու սահմաններում են, ինչը օտարալեզու խմբագրիչներն արդեն հեշտությամբ լուծում էին նախորդ դարում: Ոչ մի հայալեզու սրբագրիչ դեռևս ի վիճակի չէ լուծելու ոճական և շարահյուսական սխալ-

ներն ուղղելու խնդիրներ: Սա ունի անշուշտ իր օբյեկտիվ պատճառները: Առայժմ հայերենի լիարժեք ձևային նկարագրության հարցը լուծված չէ, հայերենի ձևայնացման տարբեր փորձեր արվել են, սակայն լիակատար ձևային նկարագրություն ասպարեզում չկա: Իսկ էլեկտրոնային սրբագրման համակարգերի հիմքում լեզվի ձևային նկարագրությունն է ընկած, հատկապես եթե ցանկանում ենք լուծել ոճական և շարահյուսական խնդիրներ:

ՀԱՅԱԼԵԶՈՒ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԵՐ

Վերջին տասնամյակներում արևելահայերենի սրբագրման համակարգեր շատերն են ստեղծել և ստեղծում: Այժմ, կարելի է ասել, արևելահայերենի սրբագրման համակարգերի առկայությունը համացանցում ապահովված է: Պետք է, սակայն, նշել, որ դրանցից որևէ մեկը լիարժեք չի սրբագրում: Այդ ծրագրերը ոչ միայն չեն ծառայում իրենց նպատակին, չեն ուղղում ուղղագրական և քերականական սխալները, այլև երբեմն անհիմն սխալների պատճառ են դառնում:

Սրբագրման առկա համակարգերի հնարավորությունները ստուգելու և թերացումները պարզելու համար ընտրել ենք գիտական տեքստից մի հատված, դրա հիմքով կազմել նմուշ-նախադասություն՝ միտումնավոր ուղղագրական սխալներով, այն մուտքագրել ենք հայերեն էլեկտրոնային սրբագրման առկա բոլոր համակարգեր և ստուգել: Պետք է նշել, որ առկա 6 համակարգերից գրեթե ոչ մեկը չլուծեց իր առջև ծա-

ռացած խնդիրը²¹: Նմուշ-նախադասությունը մուտքագրել ենք հետևյալ համակարգեր՝ *HySpell*²², *Spellchecker*²³, *Translit spellchecker*²⁴, *SpellStar*²⁵, *PowerSpell*²⁶, *Add-ons spellchecker*²⁷:

Նախքան այս համակարգերի քննությանն անցնելը ներկայացնենք, թե առհասարակ ինչպես են աշխատում էլեկտրոնային սրբագրման համակարգերը: Կիրառելու համար դրանցից մի քանիսը պետք է բեռնել համակարգիչ, ավելացնել *MS Word*, *Excell*, *PowerPoint* ծրագրային հավելվածներին և կիրառել այդտեղ, իսկ մյուսները գործում են առանձին կայքերում: Վերջինները, ի տարբերություն նախորդների, հնարավորություն են տալիս սրբագրելու ցանկացած էլեկտրոնային տեքստ անմիջապես էլեկտրոնային հարթակում: Բացի այդ՝ սրանք շահեկան են համակարգը ստեղծողների համար, քանի որ համակարգը

²¹ Հետազոտությունն իրականացրել ենք 2018-2019 թթ.: Այժմ սրբագրման նշված համակարգերում հնարավոր են որոշակի փոփոխություններ:

²² <http://www.hyspell.com/>

²³ https://www.spellchecker.net/eastern_and_western_spell_checker.html

²⁴ <http://am.translit.cc/>

²⁵ http://www.stars21.com/spelling/armenian_spell_checker.html

²⁶ <http://www.powerspell.am/>

²⁷ <https://addons.mozilla.org/en-US/firefox/addon/armenian-spell-checker-diction/>

շարունակ հարստացնելու և զարգացնելու հնարավորություն են ընձեռում: Կան նաև սրբագրման այնպիսի համակարգեր, որոնք երկու տարբերակով էլ գործում են, օրինակ՝ օտարալեզու *Grammarly*²⁸ ծրագիրը: Մեր ուսումնասիրած հայերեն էլեկտրոնային սրբագրման համակարգերից մի քանիսը ևս ունեն նման հնարավորություն, սակայն գործարկելուն առնչվող որոշակի խնդիրներ կան. դրանց կանդիդատնանք ստորև:

Ինչ վերաբերում է ուղղագրական և քերականական սխալները ուղղելուն, ապա սրբագրման համակարգերը դրանք իրականացնում են հետևյալ կերպ. ստուգումը կատարվում է շտեմարանում առկա բառամիավորների հետ համեմատելու միջոցով: Եթե ստուգվող բառը գոյություն չունի շտեմարանում, ապա երբեմն համակարգը առաջարկում է այդ բառը ներմուծել բառարան: Եթե ստուգվող բառը համակարգը սխալ է համարում, ապա այն ուղղելու համար առաջարկվում է մի քանի տարբերակ՝ ըստ ձևային նմանության: Տարբերակների քանակը սրբագրման տարբեր համակարգերում տարբեր է. հիմնականում առաջարկվում է 5-6 տարբերակ:

²⁸ <https://www.grammarly.com/>

Հայերեն էլեկտրոնային սրբագրման համակարգերից հատկապես լայն կիրառություն ունեն **Armenian PowerSpell** և **HySpell** ծրագրերը: Սրանք նաև համեմատաբար ավելի հաջող են սրբագրում:

HySpell²⁹ սրբագրման համակարգի շտեմարանն ընդգրկում է «Նոր բառգիրք հայկազեան լեզուի», է. Աղայանի «Արդի հայերենի բացատրական բառարանի» և Ստ. Մալխասյանցի «Հայերեն բացատրական բառարանի» շուրջ 158.000 բառամիավոր: Սա մի կողմից լավ է, քանի որ ընդգրկված են և՛ արևելահայերեն, և՛ արևմտահայերեն բառեր, սակայն մյուս կողմից՝ եթե առանձնացնենք արևմտահայերեն բառերը, արևելահայերենի բառերի քանակը կկրճատվի:

Դժվար չէ հասկանալ, որ փոքրաթիվ շտեմարանի առկայության դեպքում սրբագրումը լիարժեք չի լինի: Սա, թերևս, տվյալ համակարգի էական թերացումներից մեկն է: Բացի այդ՝ բեռնելուն և գործարկելուն առնչվող դժվարություններ կան: Անվճար ներբեռնելու հնարավորություն կա միայն Firefox դիտարկիչի համար, որը, պետք է նշել՝ մեր օրերում նպատակահարմար չէ, քանի որ գրեթե չի կիրառվում: Համակարգը

²⁹ <http://www.hyspell.com/>

Word 2007/2010–ին հարմարեցված ավելի հեշտ կիրառելի տարբերակներ ևս ունի. դրանք աշխատում են տարբեր օպերացիոն համակարգերով, սակայն վճարովի են, որը ևս նպատակահարմար չէ, մանավանդ որ այդ տարբերակների դեպքում էլ կիրառվում է նույն շտեմարանը՝ սահմանափակ բառաֆոնդով՝ չընդգրկելով ժամանակակից և նոր բառեր, փոխառություններ, օտար բառեր ևն:

Armenian PowerSpell³⁰ սրբագրման համակարգը ստեղծվել է MS Word, Excell, PowerPoint ծրագրերի համար: Նշենք, որ այն հեշտությամբ բեռնեցինք և կիրառեցինք Windows 7 օպերացիոն համակարգում, սակայն ոչ մի կերպ հնարավոր չեղավ կիրառել Windows 8.1-ում: Սա, բնականաբար, էական թերացում է, քանի որ ցույց է տալիս, որ այս համակարգը հնարավոր չէ կիրառել օպերացիոն համակարգերի նոր տարբերակներում, որոնք այժմ ավելի հաճախական կիրառություն ունեն: Բացի այդ՝ այս համակարգը վճարովի է, իսկ անվճար կարելի է օգտվել միայն մեկ անգամ 7 օրվա ընթացքում:

³⁰ <http://www.powerspell.am/>

Armenian PowerSpell սրբագրման համակարգի հեղինակները նշում են, որ իրենց համակարգի շտեմարանն ընդգրկում է 1 400 000 բառ: Մեր ուսումնասիրությունները ցույց տվեցին, որ շտեմարանի բառաֆոնդը հաշվելիս նկատի են առնվել ոչ թե գլխաբառերը, այլ բառաձևերը. սա նշանակում է, որ իրականում բառապաշարը սահմանափակ է: Թեև, ի տարբերություն այլ համակարգերի, այս համակարգի հիմքում առավելագույն թվով բառարաններ են, սակայն նշենք, որ նոր և նորագույն բառերը շտեմարանում չկան, լիարժեք չեն նաև օտար բառերի, հապավումների, համառոտագրությունների, հատուկ անունների դաշտերը: Այսպես, դիտարկվում է շուրջ 1000 հատուկ անուն, այն դեպքում, երբ վերջին հաշվումներով միայն աշխարհագրական անունները շուրջ 200 000 են³¹:

Այս համակարգի շտեմարանը երկու մասից է բաղկացած՝ մայր և լրացուցիչ բառարաններից: Լրացուցիչ բառարանում ընդգրկված են հատուկ անուններն ու հապավումները, որոնք բավական թերի են ներկայացված: Մայր բառարանում ներկայացված է շտեմարանի ստվար հատվածը: Ե՛վ լրացուցիչ, և՛ մայր

³¹ http://www.cadastre.am/page/ashxarhagrakan_anvanumner

բառարանը ցանկացած օգտատեր կարող է ազատ խմբագրել և մշակել: Սա, պետք է նշել, որ մասնագիտական տեսանկյունից ընդունելի չէ, քանի որ կամայականությունների և սխալների առիթ կարող է դառնալ:

Ինչպես իրենց ուղեցույցում նշում են **Armenian PowerSpell** սրբագրման համակարգի հեղինակները, այս համակարգը հնարավորություն է տալիս ստուգել և ուղղել ուղղագրական, քերականական և կետադրական սխալները: Ուղեցույցը և համակարգի նկարագրությունը կարդալուց, բազմաթիվ և՛ քերականական, և՛ ուղղագրական, և՛ կետադրական սխալներ հայտնաբերելուց հետո, թերևս, հարց է ծագում. արդյո՞ք այս ուղեցույցի հեղինակները կարող էին լիարժեք սրբագրման համակարգ ստեղծել:

Spellchecker³² և **Translit spellchecker**³³

սրբագրման համակարգերի հիմքը կազմող բառարանների և շտեմարանի բառաֆոնդի մասին տեղեկատվություն հեղինակները չեն նշում: **Spellchecker** սրբագրման համակարգը սրբագրելու երկու տարբե-

³² https://www.spellchecker.net/eastern_and_western_spell_checker.html

³³ <http://am.translit.cc/>

րակ է առաջարկում՝ հիմնական և ավելի կատարելագործված. վերջինը վճարովի է:

Մեր նմուշ-նախադասությունը ***Spellchecker*** և ***Translit spellchecker*** սրբագրման համակարգեր մուտքագրելուց հետո համակարգի՝ սխալ ճանաչած բառերի օգնությամբ հասկանալի դարձան շտեմարանի և այդ համակարգերի թերացումները: Այսպես, շտեմարանը չունի հոլովման և խոնարհման լիարժեք հարացույցներ, հատուկ անուններ, ժամանակակից տերմինաբանություն, հապավումներ ու համառոտագրություններ, օտար բառեր: Հոլովական ձևերն առաջարկելու փոխարեն առաջարկվում են անհասկանալի և չհիմնավորված տարբերակներ, ինչպես, օրինակ՝ *Աբրահամյան ին, Աբրահամյանի ն* տարբերակները:

SpellStar³⁴ սրբագրման համակարգը, թերևս, մեր ուսումնասիրած ամենաթերի համակարգն է: Այսպես, օրինակ՝ *կազմակերպուցյանը, հախորդակցման, անբիոն* սխալ ձևերը նույնիսկ կասկածելի չճանաչվեցին այս համակարգում: Ընդհանուր առմամբ պետք է նշել, որ այս ծրագիրը ոչ միայն չի ծառայում իր նպատակին, չի ուղղում ուղղագրական սխալները, ստուգ-

³⁴ http://www.stars21.com/spelling/armenian_spell_checker.html

վող բառերի համար չի առաջարկում թեքված ձևեր, բառի ձևային նմանությամբ ընտրված տարբերակներ, այլև ճիշտ բառերի համար առաջարկում է անհիմն, սխալ, ոչ մի կերպ չկիրառվող տարբերակներ՝ բառարմատից առանձնացնելով վերջավորությունը կա՛մ բացատով, կա՛մ գծիկով, ինչպես, օրինակ՝ *լեզվաբանության, լեզվաբան-ության, լեզվաբանութ յան, լեզվաբանութ-յան*:

Add-ons spellchecker³⁵ սրբագրման համակարգի հիմքում ընկած է փոքրածավալ մի շտեմարան: Բառաֆոնդը սահմանափակ է. ընդամենը պարունակում է շուրջ 150 000 բառաձև: Կիրառման տեսանկյունից այս համակարգը բոլորովին նպատակահարմար չէ: Նախ՝ շտեմարանը խիստ աղքատիկ է, հետևաբար՝ սխալները ստուգելու և ուղղելու հնարավորությունը ևս փոքր է: Բացի այդ՝ այս համակարգը նախատեսված է միայն Firefox դիտարկիչի համար, իսկ սա ամենից սակավ կիրառվող դիտարկիչն է մեր օրերում:

³⁵ <https://addons.mozilla.org/en-US/firefox/addon/armenian-spell-checker-diction/>

Ընդհանրացնելով առկա սրբագրման համակարգերի հիմնական թերությունները՝ ներկայացնենք այդ խնդիրները՝ խմբավորված:

Բառապաշարին առնչվող խնդիրներ

Սրբագրման առկա համակարգերի ընդգրկած բառապաշարի ծավալը խիստ թերի է, սահմանափակ և որևէ կերպ չի արտացոլում բառապաշարի արդի վիճակը:

1. Համակարգի հիմքում այնպիսի բառարաններ են, որոնց մեջ տեղ են գտել բազմաթիվ հնաբանություններ, հազվադեպ գործածվող բառեր, իսկ ժամանակակից, նոր, օտար բառեր այդ շտեմարանում չկան: Պետք է նշել, որ այսօր հաճախական կիրառություն ունեն հատկապես վերջինները:

2. Ծրագրում բացակայում են հատուկ անունների սրբագրման տարբերակները, հատուկ անունները չկան համակարգի շտեմարանում, չկան նաև հապավումներն ու համառոտագրությունները: Որոշ համակարգեր ընդգրկում են հատուկ անունների և հապավումների սրբագրման՝ հանրության համար բաց սրբագրման հնարավորությամբ դաշտեր, սա-

կայն սա մասնագիտորեն խիստ սխալ է, քանի որ կարող է կամայական սխալների հիմք լինել:

3. Մասնագիտական բառապաշարը, ինչպես նաև հայերենում կիրառվող օտար բառերը գրեթե ընդգրկված չեն առկա սրբագրման համակարգերում:

4. Որոշ համակարգերի շտեմարանի բառաֆոնդը հաշվելիս նկատի են առնվել ոչ թե գլխաբառերը, այլ բառաձևերը. սա նշանակում է, որ իրականում բառապաշարը սահմանափակ է: Օրինակ, եթե նկատի ունենանք, որ մեկ գոյականը եզակի և հոգնակի թվերով կարող է ունենալ 10 թեքված ձև, իսկ մեկ կանոնավոր բայը 80 դրական և 80 ժխտական, այսինքն՝ 160 թեքված ձև, ապա դժվար չէ հաշվել, որ որոշ համակարգերի շտեմարանների բառաֆոնդը բավական կրճատվում է:

Քերականական սխալների ուղղմանն առնչվող

խնդիրներ

Խիստ թերի են ուղղվում քերականական սխալները, քանի որ.

1. բացակայում են հոլովման և խոնարհման հարացույցների բոլոր ձևերը, այսինքն՝ բառաձևերը ոչ

լիարժեք են տրվում, ընդ որում՝ խիստ կամայական ընտրությամբ.

2. առաջարկվում են անհիմն, սխալ, բացարձակապես չկիրառվող տարբերակներ, ինչպես, օրինակ, բառարմատից առանձնացնելով վերջավորությունը բացատրվ կամ գծիկով: Նման գրության տարբերակներ հայերենում չկան, բացառություն են կազմում մասնագիտական տեքստերը, որտեղ վերջավորությունները կարող են գծիկով գրվել:

Գործածության սահմանափակմանն առնչվող խնդիրներ

Առկա սրբագրման համակարգերն ունեն գործածության սահմանափակում, որովհետև.

1. Անվճար ներբեռնելու հնարավորություն առաջարկվում է միայն Firefox համակարգով, որը, պետք է նշել, որ մեր օրերում նպատակահարմար չէ, քանի որ գրեթե չի կիրառվում: Բացի այդ՝ բեռնելուն և գործարկելուն առնչվող դժվարություններ կան:

2. Հնարավոր է ներբեռնել Word-ի և Outlook-ի համար նախատեսված տարբերակները, որոնք հարմա-

րեցված են Word 2007/2010 –ին, աշխատում են տարբեր օպերացիոն համակարգերով, սակայն դրանք վճարովի են, որը ևս նպատակահարմար չէ: Նշենք, սակայն, որ այս տարբերակը ևս լիարժեք չի սրբագրում:

Այսպիսով՝ ընդհանրացնելով կարելի է փաստել, որ հայերեն էլեկտրոնային սրբագրման առկա համակարգերը գրեթե չեն իրականացնում իրենց գործառույթը, հետևաբար առաջ է գալիս նոր՝ առավել լիակատար համակարգ ստեղծելու անհրաժեշտություն: Ուսումնասիրելով արևելահայերեն և օտարալեզու հնարավոր բոլոր սրբագրման համակարգերը, ռեսուրսները, կատարելով դրանց համեմատական քննություն, ինչպես նաև վեր հանելով առկա խնդիրներն ու թերությունները՝ կարելի է մշակել սրբագրման համակարգի նոր սկզբունքներ՝ առավել լիարժեք և ընդգրկուն սրբագրման համար:

Կարծում ենք՝ հայերեն էլեկտրոնային սրբագրման նոր համակարգի ստեղծումը պետք է լուծի ստորև նշվող խնդիրները.

1. Ընդգրկի էլեկտրոնային այնպիսի շտեմարան, որը լիարժեքորեն արտացոլի հայերենի արդի վիճակը, ներառի ոչ միայն հայերենի հիմնական բառաֆոնդը, այլև՝

- հատկապես վերջին շրջանում ստեղծված մեծաթիվ նոր բառեր և նորաբանություններ,
- օտարաբանություններ,
- հատուկ անուններ (աշխարհագրական անուններ, անձնանուններ, ազգանուններ),
- հապավումներ,
- համառոտագրություններ:

2. Կատարի հայերենի առավել լիակատար և ամբողջական ձևային նկարագրություն՝ կիրառելի համակարգչային տարբեր ծրագրերի համար:

3. Օգտագործելով վեբ տեխնոլոգիաների բոլոր հնարավորությունները՝ հայերենի էլեկտրոնային սրբագրման համակարգը հասանելի դառնա հանրության լայն շրջանակների համար՝ ապահովելով համակարգից անվճար օգտվելու հնարավորություն:

4. Հայերեն էլեկտրոնային սրբագրման նոր համակարգը գործակցելի դառնա օտարալեզու սրբագր-

ման այն համակարգերի հետ, որոնք առաջարկում են հայերեն տեքստերի սրբագրման հնարավորություն, ինչպես օրինակ՝ *SpellStar*³⁶, *Spellchecker*³⁷:

³⁶ http://www.stars21.com/spelling/armenian_spell_checker.html

³⁷ https://www.spellchecker.net/eastern_and_western_spell_checker.html

ARMSPELL ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԻ ԿԱԶՄՈՒԹՅԱՆ ՍԿԶՐՈՒՆՔՆԵՐԸ

Թվային տեխնոլոգիաների դարաշրջանում էլեկտրոնային տեքստը դարձել է մեր կյանքի անբաժանելի մասը: Էլեկտրոնային տեքստի առկայությունն էլ իր հերթին նպաստեց ոչ միայն էլեկտրոնային տարատեսակ ընթերցիչների ի հայտ գալուն, այլև ասպարեզ եկան էլեկտրոնային տեքստը կազմելու, խմբագրելու, սրբագրելու տարատեսակ գործիքներ: Վերջին շրջանում մեծ տարածում են գտել էլեկտրոնային սրբագրիչները: Հայերեն էլեկտրոնային առկա և գործող սրբագրիչներին, դրանց թերացումներին, առավելություններին և հեռանկարներին անդրադարձել ենք սույն ուսումնասիրության նախորդ բաժիններում և մեր հոդվածներում³⁸: Սույն բաժնում դիտար-

³⁸ Տե՛ս Մ. Սարգսյան, Հայերեն էլեկտրոնային սրբագրման առկա համակարգերի քննություն, Բանբեր Երևանի Վ. Բրյուսովի անվան պետական լեզվահասարակագիտական համալսարանի, 1(44), Երևան, 2018, էջ 598-604; **M. Sargsyan**, Some Observations on Armenian Electronic Proofreading Systems, "World Science", **№6 (34)**, Vol. 8, June 2018, Warsaw, Poland, pp. 21-24.

կենք հայալեզու տեքստերի էլեկտրոնային նոր սրբագրիչի սկզբունքներն ու առանձնահատկությունները:

«Հայերեն էլեկտրոնային սրբագրման համակարգ» գիտահետազոտական թեման իրականացվել է **երեք փուլով** երկու տարվա ընթացքում:

Առաջին փուլում (2 եռամսյակ) նախատեսված էր տվյալների շտեմարանի ու համապատասխան ծրագրաշարի ստեղծումը: Հետազոտության այս փուլի շրջանակներում կատարվել են հետևյալ աշխատանքները.

- ուսումնասիրվել են լայն տարածում գտած օտարալեզու սրբագրման համակարգերը (Grammarly, Ginger, Spell Check PlusPro, CorrectEnglish Complete և այլն)՝ տեքստը ուղղելու ժամանակ դրանցում կիրառված սկզբունքները ըստ անհարժեշտության օգտագործելու նպատակով,

- ուսումնասիրվել են հայալեզու առկա և գործող սրբագրման համակարգերը (HySpell, Spellchecker, Translit spellchecker, SpellStar, PowerSpell, Add-ons spellchecker)՝ առավելապես ուշադրություն դարձնելով այն խնդիրներին, որոնց

պատճառով դրանք ոչ բավարար մակարդակով են ուղղում տեքստը,

- առանձնացվել են տեքստերում առավել հաճախադեպ կիրառություն ունեցող ուղղագրական և քերականական սխալները՝ սրբագրման համակարգի կազմության սկզբունքները մշակելիս կիրառելու համար,

- մշակվել են հայերեն էլեկտրոնային սրբագրման նոր համակարգի կազմության սկզբունքները,

- ներկայացվել են կառուցվածքն ու ընդգրկման ծավալը, սխալներն ուղղելու եղանակը,

- համակարգվել են այն բոլոր բառարանները, որոնց նյութը ընդգրկվել է սրբագրման համակարգի հիմքում ընկած շտեմարանում (Է. Աղայանի «Արդի հայերենի բացատրական բառարան», Ա. Սուքիասյանի «Հայոց լեզվի հոմանիշների բացատրական բառարան», Ս. Գալստյանի «Դպրոցական բառակազմական բառարան», «Նոր բառեր» Ա- Դ պրակներ, Փ. Մեյթիխանյանի «Հայերենի ուղղագրական նոր բառարան», Ա. Հայրապետյանի «Օտար բառերի բառարան», Դ. Գյուրջինյանի, Ն. Հեքեքյանի «Հայերենում գործածվող տառային հապավումների բառարան», Կ. Առաքելյանի «Հատուկ անունների բառարան», Ա. Սուքիասյանի, Ք. Սուքիասյանի «Արդի հայերենի ուղղագրական-ձևա-

կազմական բառարան», ՀՀ Կառավարությանն առն-
թեր անշարժ գույքի կադաստրի կոմիտեի «Աշխար-
հագրական անվանումների հայերեն լեզվով էկզոնիմ-
ների ցանկ»),

- թվայնացվել են թղթային և էլեկտրոնա-
յին տարբերակով առկա բառարանները (թվայնաց-
ման աշխատանքներում աջակցել է ՀՀ ԳԱԱ հիմնա-
րար գիտական գրադարանի թվայնացման բաժինը),

- մշակվել է համապատասխան ծրագրա-
շարը (տեխնիկական և վեբ-կայքի աշխատանքները
կատարել է «ԱԻՍՏ ԳԼՈԲԱԼ հոլդինգը»): Վերոնշյալ
աշխատանքներն իրականացնելիս կիրառվել են
հետևյալ տեխնոլոգիաները.

- ✓ Laravel 8 framework,
- ✓ վեբ ծրագրավորման լեզուներ (PHP,
JavaScript, CSS, HTML),
- ✓ Vue.js ծրագրային ապահովում,
- ✓ MySQL բազա,
- ✓ Java ծրագրավորման լեզվով գրված որոնո-
ղական համակարգ:

Երկրորդ փուլում (4 եռամսյակ) նախատեսված
էր շտեմարանում ընդգրկվելիք բառապաշարի ձևային

նկարագրություն: Հետազոտության այս փուլի շրջանակներում կատարվել են հետևյալ աշխատանքները.

- մշակվել է վերը նշված բոլոր բառարանների բառարանային նյութը՝ առանձնացնելով միայն գլխաբառերը,

- կատարվել է ընդգրկվելիք բառապաշարի ձևային նկարագրություն (շուրջ 300 000 գլխաբառ՝ դրանցից յուրաքանչյուր գոյականի համար 10 թեքված ձև, դերանվան համար՝ 12, բայի համար՝ 160, ըստ անհրաժեշտության նշված են նաև զուգաձևությունները),

- կազմվել են ձևաբանական փոփոխությունների ենթարկվող բառերի հոլովման և խոնարհման հարացույցները, մշակվել է ձևային նկարագրության յուրատեսակ ալգորիթմ:

Երրորդ փուլում (2 եռամսյակ) նախատեսված էր համակարգի հնարավոր սխալների ու թերությունների շտկում: Հետազոտության այս փուլի շրջանակներում կատարվել են հետևյալ աշխատանքները.

- Խմբագրվել և սրբագրվել են էլեկտրոնային համակարգի կողմից ավտոմատ սերված (գեներացված) բառերը,

- փորձարկվել է շտեմարանի որոնողական աշխատանքը, կատարվել են բարելավման աշխատանքներ:

«Հայերեն էլեկտրոնային սրբագրման համակարգ» գիտահետազոտական թեմայի շրջանակներում, ինչպես նշվել է վերևում, ստեղծվել է սրբագրման այնպիսի համակարգ, որի միջոցով փորձ է արվում ամբողջական և լիարժեք կերպով սրբագրել արդի արևելահայերենով գրված ցանկացած տեքստ, ինչպես նաև օգտագործելով վեբ տեխնոլոգիաների բոլոր հնարավորությունները, հայերեն էլեկտրոնային սրբագրման համակարգը հասանելի դարձնել հանրության լայն շրջանակներին՝ ապահովելով համակարգից հեշտությամբ օգտվելու հնարավորություն:

ARMSPELL ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԻ ԱՌԱՆՁՆԱՀԱՏԿՈՒԹՅՈՒՆՆԵՐԸ

Հանրության լայն շրջանակներին հասանելի դարձնելու նպատակով **ARMSPELL** էլեկտրոնային սրբագրման համակարգը գործում է համացանցային առանձին կայքում՝ **armspell.am** վեբ հասցեի ներքո՝ գրանցված որոնողական բոլոր համակարգերում: Կիրառման տեսակետից այն պարզ է և արագագործ, քանի որ դրանից օգտվելու համար որևէ ծրագիր ներբեռնելու անհրաժեշտություն չկա:

Տվյալների շտեմարանի կառուցվածքը/ կադապարը հնարավորություն է տալիս ոչ միայն նախապես մուտքագրել տվյալները, այլև համալրել հետագայում:

Մշակվել է նաև համապատասխան ծրագրաշարը: Համակարգն ունի կառավարման վահանակ (այսուհետ՝ ԿՎ), որը ծառայում է համակարգի շտեմարանում նոր բառեր ավելացնելու, ինչպես նաև վրիպակները, հնարավոր սխալները շտկելու համար: ԿՎ-ը բաղկացած է մի քանի էջից ու գործիքներից: Գործիք-

ներից մեկը նախատեսված է գլխաբառերը մուտքագրելու համար, որի համար օգտագործվում է հատուկ տեխնոլոգիա, որը հնարավորություն է տալիս շտեմարանում բառեր մուտքագրել և՛ մեծ խմբաքանակով, և՛ առանձին-առանձին: Գործիքներից մյուսը նախատեսված է արդեն մուտքագրված բառերն ըստ խոսքի մասերի դասակարգելու համար: Այս էջում առկա է նաև գլխաբառերի հնարավոր թեքված տարբերակները ավտոմատ գեներացնելու հնարավորություն, այնուհետև՝ թե՛ մուտքագրելու, թե՛ խմբագրելու, թե՛ հեռացնելու հնարավորություն:

Կայքի օգտատերերը սրբագրման համակարգը կիրառելիս որոնման դաշտում կարող են մուտքագրել ստուգվող տեքստը և ստուգել այդ տեքստի բառերի ուղղագրական և քերականական սխալները: Սխալ գտնելու դեպքում համակարգը սխալն ուղղելու մի քանի տարբերակ է առաջարկում:

Համակարգը ուղղագրական և քերականական սխալների ստուգումը կատարում է շտեմարանում առկա բառերի հետ համեմատելու միջոցով: Եթե ստուգվող բառը բացակայում է շտեմարանից, ապա այն համակարգի կողմից համարվում է ուղղագրական սխալ

ունեցող բառ: Միսալ ունեցող բառը գտնելուց հետո համակարգը համեմատում է այն շտեմարանում առկա բառերի հետ և գտնում ձևային առումով ամենամոտ բառերը՝ ներկայացնելով այն օգտատիրոջը՝ որպես հնարավոր ճիշտ տարբերակ:

Էլեկտրոնային նոր սրբագրիչի շտեմարանը լիարժեք կերպով արտացոլում է հայերենի արդի վիճակը, քանի որ ներառում է ոչ միայն հայերենի հիմնական բառաֆոնդը, այլև՝

- հատկապես վերջին շրջանում ստեղծված մեծաթիվ նոր բառեր և նորաբանություններ,
- օտարաբանություններ,
- հատուկ անուններ (աշխարհագրական անուններ, անձնանուններ, ազգանուններ),
- հապավումներ,
- համառոտագրություններ:

Կայքից օգտվողների թիվը որևէ սահմանափակում չունի: Ծառայելով հասարակության լայն շրջանակներին՝ այն **կարող է օգտակար լինել**.

ա. արևելահայերի համար ցանկացած տեքստի սրբագրման ժամանակ (գիտական, հրապարակախոսական, գեղարվեստական և այլն),

բ. արևմտահայերի համար, որոնք, ուղղագրական համակարգերի տարբերություններով պայմանավորված, ոչ միշտ են կարողանում ինչպես ճիշտ ուղղագրությամբ արևելահայերեն տեքստեր գրել, այնպես էլ առհասարակ լիարժեք կերպով կիրառել հայերենի քերականական կանոնները:

գ. օտարերկրացիների համար, որոնք բավարար լեզվական գիտելիքներ չունենալու պատճառով չեն կարող կառուցել իրենց գրավոր խոսքը պատշաճ մակարդակով:

Այսպիսով՝ կայքը և էլեկտրոնային նոր սրբագրիչը կարող է կիրառել ցանկացած մարդ՝ անկախ սեռից, տարիքից, կրթական մակարդակից, բնակության վայրից:

Ծրագրաշարը մշակված է այնպես, որ հետագայում հնարավոր է ընդլայնել համակարգի հնարավորությունը՝ ընդգրկելով շարահյուսական, կետադրական և ոճական սխալների ուղղումը: Այն կարելի է

գործակցելի դարձնել նաև օտարալեզու սրբագրման համակարգերի հետ, որոնց միջոցով կարելի է կատարել հայերեն տեքստերի սրբագրում (http://www.stars21.com/spelling/armenian_spell_checker.html, https://www.spellchecker.net/eastern_and_western_spell_checker.html, <https://addons.mozilla.org/en-US/firefox/addon/armenian-spell-checker-diction/>):

Էլեկտրոնային նոր սրբագրիչի 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 գործառությանը: Այն մեծ հեռանկարներ է բացում ինչպես տեսական-լեզվաբանական, այնպես էլ գործնական-կիրառական խնդիրների լուծման համար:

**ARMSPELL ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ
ՀԱՄԱԿԱՐԳԻ ՇՏԵՄԱՐԱՆԻ
ԱՌԱՆՁՆԱՀԱՏԿՈՒԹՅՈՒՆՆԵՐԸ**

Սրբագրման ցանկացած համակարգի հաջող գործարկման հիմքում ոչ միայն հմտորեն մշակված ծրագրաշարն է, այլև տվյալների հարուստ շտեմարանը: Բացի այդ՝ սրբագրիչի տվյալների շտեմարանում ընդգրկվելիք բառապաշարը պետք է ձևային նկարագրության ենթարկել: Լեզվի ձևայնացման գործընթացը բարդ և պատասխանատու աշխատանք է, ի հայտ են գալիս բազմաթիվ խնդիրներ՝ կապված ոչ միայն ձևայնացման սկզբունքների մշակման ճշգրտության հետ, այլև լեզվի կառուցվածքային տիպի, ձևաբանական փոփոխությունների ենթարկվող բառերի հոլովման և խոնարհման հարացույցների առանձնահատկությունների հետ և այլն: Լեզվի ձևային նկարագրության սկզբունքները մշակելիս պետք է հաշվի առնել տվյալ լեզվի կառուցվածքային առանձնահատկությունները: Ձևայնացման լիարժեքությունն ապահովելու համար պակաս կարևոր չէ նաև այն նպատակի հստակ սահմանումը, որին ծառայելու է

ձևայնացումը: Հաշվի առնելով ժամանակակից հայերենի ձևայնացման առանձնահատկությունները և դրանցով պայմանավորված դժվարությունները՝ սույն բաժնում քննության ենք ենթարկում այս բազմազանությունը ապահովող ձևային նկարագրության խնդրահարույց հարցերը, ներկայացնում դրանց լուծման ուղիները, սկզբունքները, որոնք ընկած են մեր մշակած էլեկտրոնային սրբագրիչի տվյալների շտեմարանի ստեղծման հիմքում:

Ժամանակակից հայերենի ձևային նկարագրությունը հանգամանալից աշխատանք է պահանջում: Խնդրահարույց հարցեր են առաջացնում մեր լեզվի հատկապես հետևյալ առանձնահատկությունները.

1. Ժամանակակից հայերենին բնորոշ են *քերականական կարգերի առաջությունը*, ինչպես նաև դրանց արտահայտման միջոցների բազմազանությունը: Ինչպես գիտենք, հայերենում միևնույն քերականական իմաստը կարող է արտահայտվել մի քանի տարբերակային ձևերով:

2. Հոլովման և խոնարհման հարացույցներում, նորմատիվ ձևերից բացի, առկա են *անկանոնություն-*

ներ, որոնք շեղվում են հոլովման և խոնարհման ընդհանուր օրինաչափություններից: Դրանց համար անհրաժեշտ է հստակ սահմանումներ մշակել:

3. Ժամանակակից հայերենը գերազանցապես կցական լեզու է, սակայն և՛ խոնարհման, և՛ հոլովման հարացույցներում առկա են *անջատականության և թեքականության դրսևորումներ*, որը ձևայնացման գործընթացը որոշակիորեն դժվարացնում է:

Դժվար չէ նկատել, որ վերոնշյալ հանգամանքները մեծացնում են ձևային նկարագրության ծավալը՝ ստեղծելով տարբերակային ձևերի առատություն:

Բառապաշարի ձևային նկարագրությունը հնարավորություն է տալիս բացահայտելու ինչպես կանոնավոր կազմությունների, այնպես էլ բոլոր տարբերակային ձևերի, շեղումների ու անկանոնությունների ամբողջական պատկերը՝ հաճախականության տվյալներով և դրանց ավտոմատ վերլուծության ու սերման հնարավորություններով: Այսինքն՝ էլեկտրոնային շտեմարանը ներկայացնում է արդի հայերենի բառապաշարի բառամիավորների ձևաբանական բնութագիրը:

Հայերենի բառապաշարը ձևայնացնելու և համապատասխան էլեկտրոնային շտեմարան ստեղծելու համար լուծվել են հետևյալ **խնդիրները**.

1. ուսումնասիրելով արդի լեզվաբանության մեջ ընդունված լեզվի ձևային նկարագրության սկզբունքներն ու եղանակները՝ մշակվել են հայերենի բառապաշարի ձևային նկարագրության ընդհանրական սկզբունքներ,

2. որոշվել է բառապաշարի մոտավոր ծավալը, որի ձևային նկարագրությունը նախատեսվում է իրականացնել (այժմ շտեմարանում մուտքագրված է շուրջ 300 000 գլխաբառ, դրանցից յուրաքանչյուր գոյականի համար 10 թեքված ձև, դերանվան համար՝ 12, բայի համար՝ 160, ըստ անհրաժեշտության նշված են նաև զուգաձևությունները),

3. բառարաններից առանձնացվել և մշակվել են համապատասխան բառամիավորները (շուրջ 300 000 գլխաբառ),

4. կատարվել է ընտրված բառամիավորների ձևային նկարագրություն ձևաբանական տեսակետից,

5. կատարվել է վերոնշյալ բառերի ձևաբանական վերլուծություն,

6. առանձնացնելով լեզվական նյութը՝ մշակվել են համապատասխան էլեկտրոնային շտեմարանի կազմության սկզբունքները,

7. ստեղծվել է էլեկտրոնային շտեմարանի կառուցվածքը/կադապարը:

Համակարգն ունի կառավարման վահանակ (այսուհետ՝ ԿՎ), որը ծառայում է համակարգի շտեմարանում նոր բառեր ավելացնելու, ինչպես նաև վրիպակները, հնարավոր սխալները շտկելու համար: ԿՎ-ը բաղկացած է մի քանի էջերից: Առաջին էջը նախատեսված է բառերի մուտքագրման համար, որի համար օգտագործվում է հատուկ տեխնոլոգիա, որը հնարավորություն է տալիս շտեմարանում բառեր մուտքագրել և՛ մեծ խմբաքանակով, և՛ առանձին-առանձին: Երկրորդ էջը նախատեսված է արդեն մուտքագրված բառերն ըստ խոսքի մասերի դասակարգելու համար: Այս էջում առկա է նաև գլխաբառերի հնարավոր թեքված տարբերակները թե՛ մուտքագրելու, թե՛ խմբագրելու, թե՛ հեռացնելու հնարավորություն:

Համակարգը ուղղագրական և քերականական սխալների ստուգումը կատարում է շտեմարանում առ-

կա բառերի հետ համեմատելու միջոցով: Բառերի համեմատությունը կատարվում է տեքստերի նմանության Լևենշտեյնյան հեռավորության չափման միավորի հիման վրա (Levenshtein distance) ստեղծված նոր ալգորիթմով:

Հայերենի էլեկտրոնային սրբագրիչի տվյալների շտեմարանը կազմված է հետևյալ աղյուսակներից՝ **բառույթ, խոսքի մաս, հոգնակի թվի տիպ, հոգնակի թիվ, թեքման տիպ, հոգնակի թվի թեքման տիպ**:

Բառույթ -բառի ուղիղ ծնն է, մուտքագրվում է շտեմարան թե՛ առանձին, թե՛ խմբաքանակով, մուտքագրելիս կրկնությունները ավտոմատ կերպով հեռացվում են,

Խոսքի մաս- մեկից ավելի ընտրելու հնարավորությամբ ընտրովի դաշտ է՝ գոյական, ածական, թվական, դերանուն, բայ, մակբայ, կապ, շաղկապ, վերաբերականներ, ձայնարկություններ, կարող է մնալ չլրացված,

Հոգնակի թվի տիպ- մուտքագրվում է համապատասխան տիպի պիտակը, որի հիման վրա ավտոմատ կազմվում և լրացվում է **Հոգնակի թիվ** դաշտը,

Հոգնակի թիվ – կարող է չլրացվել,

Թեքման տիպ – մուտքագրվում է համապատասխան տիպի պիտակը, որի հիման վրա ավտոմատ կազմվում և լրացվում են թեքված ձևերը՝ ամեն մեկը իր տիպին բնորոշ դաշտերի քանակով,

Հոգնակի թվի թեքման տիպ – մուտքագրվում է համապատասխան տիպի պիտակը, որի հիման վրա ավտոմատ կազմվում և լրացվում են հոգնակի թվով թեքված ձևերը՝ ամեն մեկը իր տիպին բնորոշ դաշտերի քանակով:

Համակարգը ավտոմատ կազմում է մուտքագրված բառերի թեքված ձևերը՝ ըստ մուտքագրված պիտակի: Օրինակելի նմուշները պատրաստի կազմություններ են, որոնք նախապես մուտքագրվում են համակարգ: Կարող են լինել նույն բառույթի մի քանի տիպեր: Երկրորդ, երրորդ և ավել տիպերը կազմելու համար **Թեքման տիպ** դաշտում ավելացվում է նոր պիտակ մուտքագրելու դաշտ, որը մուտքագրելուց հետո կազմվում են համապատասխան ձևերը:

Բոլոր դաշտերը կարող են խմբագրվել: Ավտոմատ կազմված ձևերը կարելի է խմբագրել՝ պահպա-

նելով նոր ձևը կամ ուղղելով համակարգի սխալ կազմած ձևը, դաշտերը կարող են մնալ նաև չլրացված կամ դատարկ:

Օգտատիրոջ մուտքագրած տեքստը սրբագրելիս համակարգը որոնում և համապատասխանեցում է անում ոչ միայն նախապես շտեմարան մուտքագրված **Բառույթ** դաշտի հետ, այլև տիպային պիտակերի հիման վրա կազմված/ գեներացված համապատասխան ձևերի հետ:

Թեքման տիպի պիտակները տեղեկություն են հաղորդում տվյալ բառույթի խոսքիմասային պատկանելության, թեքման տիպի, դրա ենթատիպի մասին:

Եզակի թվով բառույթի հոգնակի թիվը կազմելու համար **Հոգնակի թիվ** դաշտին կից **Հոգնակի թվի տիպ** դաշտում (կարող է ավելացվել նոր պիտակ մուտքագրելու դաշտ) մուտքագրվում է հոգնակի թվի ուղիղ/եղակետային ձևը կազմելու պիտակը: Այդ պիտակով կազմվում է տվյալ բառույթի հոգնակի ձևը:

Հոգնակի թվի թեքված ձևերը կազմելու համար **Հոգնակի թվի թեքման տիպ** դաշտում լրացվում է համապատասխան պիտակը, որի հիման վրա կազմ-

վում են համապատասխան հոգնակի ձևերը: Դաշտը կարող է մնալ չլրացված:

Ըստ վերոնշյալ սկզբունքների՝ միայն գոյականը եզակի թվում ունի տասնմեկ տիպ, տիպերի հիմնական մասն ունեն ենթատիպեր, ենթատիպերի ընդհանուր թիվը շուրջ հարյուր է: Նկատի են առնվել նաև անկանոն ձևերը:

Այսպիսով՝ հաշվի առնելով ժամանակակից հայերենի ձևայնացման առանձնահատկությունները և մշակելով դրանք նկարագրելու ուրույն համակարգ՝ կարող ենք փաստել, որ հայերեն էլեկտրոնային նոր սրբագրիչի շտեմարանը լիարժեք կերպով արտացոլում է հայերենի արդի վիճակը, քանի որ ներառում է ոչ միայն հայերենի հիմնական բառաֆոնդը, այլև թեքվող ձևերի տարբերակային հնարավոր բոլոր ձևերը:

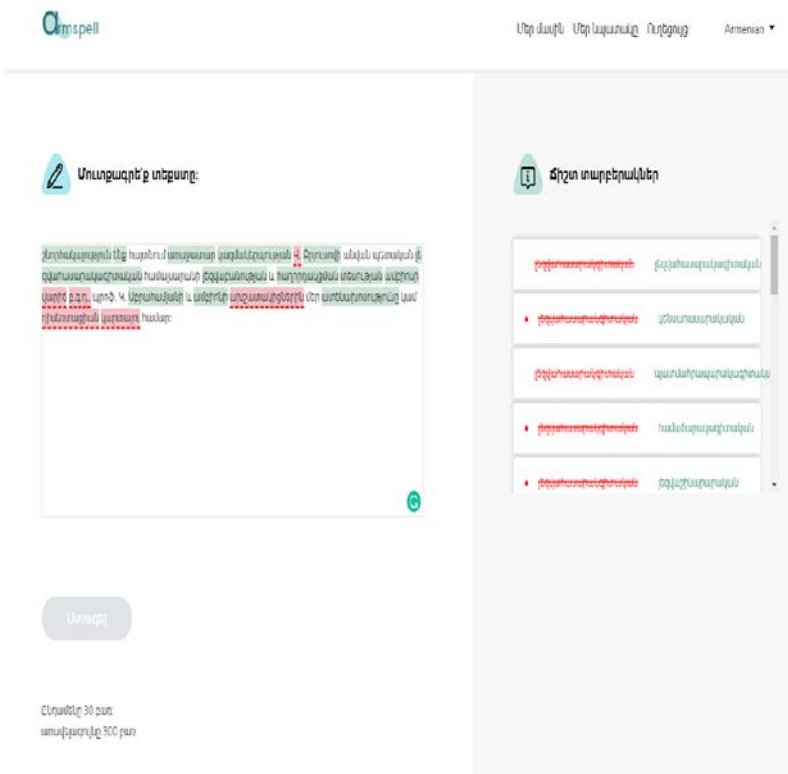
ARMSPELL ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ ՀԱՄԱԿԱՐԳԻ ԿԻՐԱՌՈՒԹՅԱՆ ՈՒՂԵՑՈՒՅՑ

Հանրության լայն շրջանակներին հասանելի դարձնելու նպատակով **ArmSpell** հայերեն էլեկտրոնային սրբագրման համակարգը գործում է համացանցային առանձին կայքում՝ **armspell.am** վեբ հասցեի ներքո, որը գրանցված է որոնողական բոլոր համակարգերում³⁹:

Կիրառման տեսակետից համակարգը պարզ է և արագագործ, քանի որ դրանից օգտվելու համար որևէ ծրագիր ներբեռնելու անհրաժեշտություն չկա: Այսինքն՝ եթե օգտատերը ցանկանում է օգտվել **ArmSpell սրբագրիչից**, անհրաժեշտ է պարզապես համացանցի հասանելիություն:

³⁹ «Հայերեն էլեկտրոնային սրբագրման համակարգ», <https://armspell.am/hy>

Հայերեն էլեկտրոնային սրբագրման համակարգ



Հիշատակելի է, որ սրբագրման այս համակարգը հասանելի է ամենքին բաց մատչելիությամբ, ինչը նշանակում է, որ կարիք չկա նախապես գրանցվելու, համակարգ մուտք գործելու, ծրագիր ներբեռնելու,

ծրագրի նորացնելու, ծրագրի նորացված տարբերակը ներբեռնելու և այլն:

ArmSpell հայերեն էլեկտրոնային սրբագրման համակարգը միալեզվյան սրբագրիչ է, հայտնաբերում և ուղղում է արդի արևելահայերենով գրված բառերում և տեքստերում առկա ուղղագրական և ձևաբանական սխալները:

Սրբագրման համակարգը նախատեսված է լայն հանրության՝ թե՛ հայաստանաբնակ, թե՛ սփյուռքահայ, թե՛ օտարերկրացի օգտատերերի համար, հետևաբար՝ կայքը երկլեզու է, այն կարելի է գործարկել հայերեն և անգլերեն:

Մեր մասին Մեր նպատակը Ուղեցույց

Armenian ▼



Ճիշտ տարբերակներ

Կայքն ունի երեք բաժին՝ «Մեր մասին», «Մեր նպատակը», «Ուղեցույց», որոնք ներկայացնում են **ArmSpell** հայերեն էլեկտրոնային սրբագրման համակարգը, ֆինանսավորման և ղեկավարի տվյալները, նախագծի նպատակը, ինչպես նաև համակարգից օգտվելու համառոտ ուղեցույցը:

1



armspell.am/hy



1. **ArmSpell** հայերեն էլեկտրոնային սրբագրման համակարգից օգտվելու համար

անհրաժեշտ է դիտարկչի որոնումների տողում մուտքագրել սրբագրիչի վեբ հասցեն՝ <https://armspell.am/hy>:

2. **ArmSpell** հայերեն էլեկտրոնային սրբագրման համակարգ մուտք գործելուց հետո էկրանին բացվում է պատուհան, որտեղ պետք է մուտքագրել ստուգման ենթակա բառը, բառերը կամ տեքստը:

2

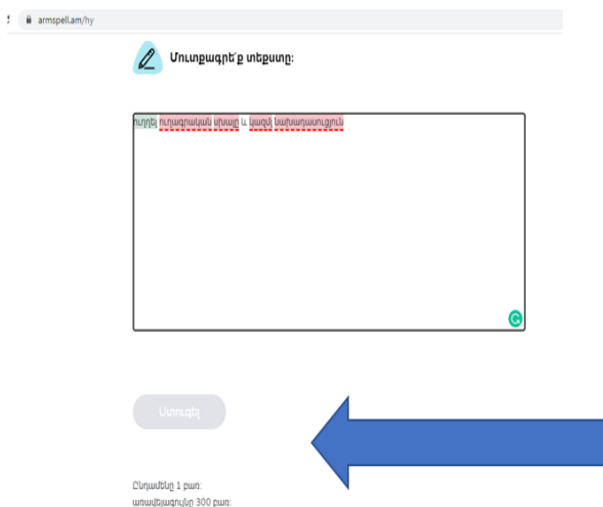


Մուտքագրե՛ք տեքստը:



ArmSpell կայքի օգտատերերը սրբագրման համակարգը կիրառելիս որոնման դաշտում կարող են մուտքագրել առավելագույնը 300 բառ և ստուգել այդ բառերի ուղղագրական և քերականական սխալները:

3



The screenshot shows the ArmSpell website interface. At the top, there is a search bar with the URL 'armspell.am/hy'. Below the search bar, there is a blue pencil icon and the text 'Մուտքագրե՛ք տեքստը:' (Enter the text). Below this, there is a large white rectangular box for text input. At the bottom of the box, there is a small green circular icon. Below the input box, there is a blue button with the text 'Ստուգել' (Check). To the right of the button, there is a large blue arrow pointing left. Below the button, there is a small text label: 'Օրդամեկ 1 բառ: առավելագույնը 300 բառ:' (Order 1 word: maximum 300 words).

3. Բառը, բառերը կամ տեքստը մուտքագրելուց հետո անհրաժեշտ է սեղմել «Ստուգել» կոճակը, սպասել մի քանի վայրկյան, մինչև համակարգը կհայտնաբերի լեզվական սխալները:

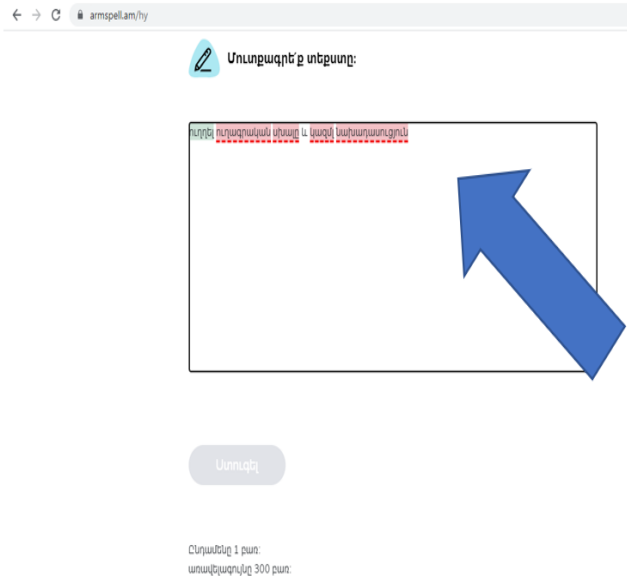
Համակարգը ուղղագրական և քերականական սխալների ստուգումը կատարում է սրբագրման համակարգի հիմքում ընկած շտեմարանում առկա բառերի հետ համեմատելու միջոցով:

Օգտատիրոջ մուտքագրած տեքստը սրբագրելիս համակարգը որոնում և համապատասխանեցում է անում ոչ միայն նախապես շտեմարան մուտքագրված **Բառույթ** դաշտի հետ, այլև տիպային պիտակերի հիման վրա կազմված/գեներացված համապատասխան ձևերի հետ: Այսինքն՝ ուղղումներ են առաջարկվում նաև տվյալ բառի թեքված տարբերակներից:

Եթե ստուգվող բառը բացակայում է շտեմարանից, ապա այն համակարգի կողմից համարվում է ուղղագրական սխալ ունեցող բառ:

4. Լեզվական սխալներ հայտնաբերելուց հետո համակարգը օգտատիրոջ մուտքագրած տեքստում սխալ ունեցող բառերն ընդգծում է կարմիր երանգով:

4



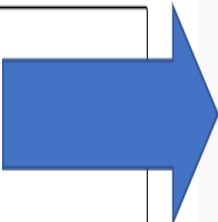
5. Սեղմելով կարմիր երանգով ընդգծված սխալ բառերի վրա՝ էկրանի աջ կողմում բացվում է պատուհան, որտեղ երևում են սխալն ուղղելու ճիշտ տարբերակները:

5



Մուտքագրե՞ք տեքստը:

Իսկով: Խոլաղոսական տեսող և կազմ նախադասություն



Ճիշտ տարբերակներ

- նախադասություն նախադասություն
- նախադասություն համարասանություն
- նախադասություն նախադասություն
- նախադասություն նախադասություն
- նախադասություն բառարանություն

համակարգը համեմատում է այն շտեմարանում առկա բառերի հետ և գտնում ձևային առումով ամենամոտ

բառերը՝ ներկայացնելով այն օգտատիրոջը՝ որպես հնարավոր ճիշտ տարբերակներ:

Հիմնվելով բառի ձևային նմանության վրա՝ համակարգը սխալ գտնելու դեպքում այն ուղղելու 6-10 տարբերակ է առաջարկում:

**Ճիշտ տարբերակներ**

- նախադասություն նախադասություն
- նախադասություն համադասություն
- նախադասություն նախագահություն
- նախադասություն նախատեսություն
- նախադասություն բառադասություն

Օրինակ՝ **նախադասություն** սխալ բառն ուղղելու համար համակարգը որոնել և ըստ նմանության գտել ու առաջարկում է հետևյալ բառերը՝ **նախադասություն** (տվյալ դեպքում ճիշտ տարբերակն է), **համադասություն**, **նախագահություն**, **նախադասություն**, **բառադասություն** ևն:

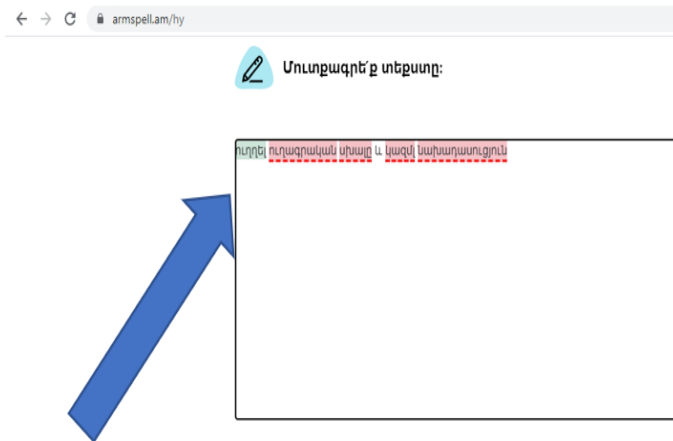
6. Համակարգի կողմից ձևավորված ճիշտ տարբերակների մեջ տեսնելով նախընտրելի տարբերակը՝ օգտատերը սեղմում է այդ ճիշտ տարբերակի վրա:

6



7. Սեղմելով համակարգի առաջարկած ճիշտ տարբերակներից մեկի վրա՝ օգտագիրոջ մուտքագրած տեքստում կարմիր երանգավորում ստացած սխալ բառը փոխարինվում է այդ ճիշտ տարբերակով և ստանում կանաչ երանգավորում:

7



Յուրաքանչյուր սխալ բառի դեպքում պետք է կատարել նույն գործողությունները և ուղղել սխալը:

Ինչպես նշել ենք վերևում, եթե որևէ բառ կամ բառաձև ընդգրկված չէ շտեմարանում, ապա համակարգի կողմից համարվելու է ուղղագրական կամ քերականական սխալ ունեցող բառ:

Այսպես, օրինակ՝ եթե օգտատերը մուտքագրի որևէ օտար բառ, որը գործածական չէ հայերենում կամ գրանցված չէ բառարաններում (օտար բառերի առկա բառարանների բառանյութը ամբողջությամբ ընդգրկված է սրբագրիչի հիմքում ընկած շտեմարանում), համակարգը չի կարող գտնել այն շտեմարանում: Հետևաբար՝ նմանատիպ բառերը ստուգելիս համակարգը ճիշտ տարբերակ չի կարող առաջարկել:

Ինչ վերաբերում է **ArmSpell** հայերեն էլեկտրոնային սրբագրման համակարգի շտեմարանին առհասարակ, ապա պետք է նշել, որ այն բավական ընդգրկուն է, ներառում է ժամանակակից արևելահայերենի առկա բառարանների բառարանային նյութը և ունի մի կարևոր առավելություն. այն կայուն չէ, շարունակ համալրվում է, այսինքն՝ օգտատերերը, կիրառելով սրբագրման համակարգը, բաց մատչելիությամբ կարող են օգտվել

համակարգի թարմացումներից և նորացված շտեմարանից:

ArmSpell հայերեն էլեկտրոնային սրբագրման համակարգի կայքն ունի «*Հեղադարձ կապ*» բաժին, որի միջոցով կայքի օգտատերերը կարող են իրենց դիտարկումներն ու առաջարկությունները ուղարկել սրբագրիչը մշակողներին և նպաստել սրբագրիչի հետագա մշակմանը և բարելավման աշխատանքներին:

SUMMARY

Although a small number of Armenian Spell Checkers have been created recently, they fail to correct even the most frequent mistakes in texts. So we can surely say that the quality of the Armenian Spell Checkers is low. It refers to both the volume of the vocabulary included in those programs and grammar, as not all types of grammatical errors can be corrected by them.

All Armenian Spell checkers have been created only by programmers but not linguists. This may be considered the first shortcoming of those programs as there is no guarantee that the corrections made by those programs are right.

Summarizing other major shortcomings of the existing Spell Checkers, let's mention some of them.

- **The vocabulary** of those programs is extremely incomplete, limited and does not reflect the current state of the Armenian vocabulary.

1. The programs are based on dictionaries that include many archaisms, rarely used words, but there aren't any neologisms and barbarisms in the database. It should be noted that today the latter is frequently used.

2. The programs do not have proofreading options for proper names, as their database does not include proper names, abbreviations and acronyms.

3. Some proofreading systems include open source public proofreading fields, but this is a mistake from professional perspective because it can become a basis for arbitrary mistakes.

4. Widely used terms practically are not included in the existing spell checkers.

5. Calculating the database vocabulary of some programs, it has been taken into account not headwords, but morphological forms. This clearly shows that the vocabulary of those programs is actually limited. For example, if a noun with singular and plural numbers can have 10 morphological forms and a verb 160, it is

not difficult to count that the vocabulary is sufficiently reduced.

- **Grammatical errors** are inaccurately corrected because those programs:

1. don't include all forms of the systems of declension and conjugation in the database,

2. suggest unreasonable, incorrectly-used variants, such as separating an ending from a word-root with spatial or a dash.

- The available spell checkers have **limited usage** because:

1. To use those programs, you have to buy a full package of the software, add in Microsoft Word, Excel, PowerPoint text editors and use them in your PC. This implies that only a limited number of people can use them. Moreover, even the creators can't later modify or replenish their own software and correct errors made in those programs.

2. Free download option is offered only by Firefox browser, while the majority of Internet users use Google and Yandex search engines. This means they simply can't have access to this text proofreading app. In

addition, it can only be used to correct errors of the texts sent via the Internet (emails, messages, publications on social networking sites).

However, it can't be used to correct text written by text editors.

The aim of "The Armenian Proofreading System" is to create a proofreading system which allows all Armenians around the world proofread texts written in the modern Eastern Armenian language.

To reach this aim **the following problems** have been underlined and resolved:

1. have been researched reputable, widely used foreign-language spell and grammar checkers. Some of the key principles used in those programs have been applied to the Armenian Electronic Proofreading System.
2. have been examined all the Armenian existing spell checkers, focusing on the main reasons why they make corrections at such a low level or fail to detect and correct some kind of mistakes,

3. have been developed the principles of the new electronic proofreading system, to present its structure and volume of inclusion, error correction method, etc.

4. have been systematized all the dictionaries which material should be included in the database of the following proofreading system;

5. have been digitized all existing dictionaries in paper and electronic form,

6. has been elaborated the vocabulary of all dictionaries by specifying only the headwords,

7. has been made a formal description of the vocabulary including the morphological changes of the words, i.e. paradigms of the declension and conjugation,

8. as a final result, has been presented a new and more complete system of Armenian electronic proofreading,

9. using all the possibilities of web technologies has been made the Armenian electronic proofreading system accessible to the public, providing free access to the system.

In order to make the project accessible to the general public, it is available on its own website under the website **armspell.am**, registered in all search

engines⁴⁰. The usage of the Proofreading system is very convenient, simple and fast for users, as it is not necessary to download any other program.

The project has been planned and finished in **three stages**:

It was intended for the *first phase* (2 quarters) to create the database and corresponding software. At this stage of the project, the structure of the database has been created and finalized, the data has been entered and only after all these steps the appropriate software has been developed based on the database.

In the second phase (4 quarters) a formal description of the vocabulary to be included in the database has been provided, separating headwords, declined forms and morphological alterations, etc.

In the third phase (2 quarters) it has been planned to correct possible mistakes and shortcomings of the system, and it has done successfully.

⁴⁰ “Armenian Electronic Proofreading System”, <https://armspell.am/hy>

As **a final result**, we have:

1. "The Armenian Electronic Proofreading System", with a free access to the system and available for the wide range of people.

2. An electronic database including around 300,000 headwords that fully presents the current state of the Armenian language, as it includes not only the main vocabulary of Armenian, but also

- a number of new words and neologisms recently created,
- barbarisms,
- proper names (geographical names, personal names, surnames),
- acronyms,
- abbreviations.

3. A complete formal description of the Armenian language, which later can successfully be applied to different computer programs.

4. Articles and a study devoted to the analysis of the collected material, reports that contain the information on the principles of creating the Armenian Electronic Proofreading System and the volume of inclusion, its application, possible replenishment, etc.

5. Digitized new dictionaries.

The number of the users of the project is not limited. This project **can serve** a wide range of people, in particular, it can be useful to:

1. Eastern Armenians to proofread any text (scientific, publicist, artistic, etc.) written in Eastern Armenian,
2. Western Armenians who are not always able to write Eastern Armenian texts correctly, and fully use the grammatical rules of the Armenian language, due to differences in spelling systems.
3. Foreigners who can not write properly due to the lack of linguistic knowledge.

The following study may be useful for anyone both in Armenia and in the Diaspora, regardless of gender, age, educational level or place of residence.

“The Armenian Electronic Proofreading System” has a great perspective as it can be made collaborative with foreign language proofing systems that offer Armenian text correction opportunities

(http://www.stars21.com/spelling/armenian_spell_checker.html,
https://www.spellchecker.net/eastern_and_western_spell_checker.html, <https://addons.mozilla.org/en-US/firefox/addon/armenian-spell-checker-diction/>).

The Armenian Electronic Proofreading System will greatly contribute to the use of computer technology in linguistic research, elimination of language errors in texts through computer programs. In the broad sense, the Armenian language will become a web language.

ՕԳՏԱԳՈՐԾՎԱԾ ԳՐԱԿԱՆՈՒԹՅԱՆ ՑԱՆԿ

1. Սարգսյան Մ. Ա., Հայերեն էլեկտրոնային սրբագրման առկա համակարգերի քննություն, Բանբեր Երևանի Վ. Բոյուսովի անվան պետական լեզվահասարակագիտական համալսարանի, 1(44), Երևան, 2018, էջ 598-604:
2. Սարգսյան Մ. Ա., Էլեկտրոնային սրբագրիչի տվյալների շտեմարանի բառապաշարի ձևային նկարագրության խնդիրների մասին, «Բանբեր Երևանի Վ. Բոյուսովի անվան պետական լեզվահասարակագիտական համալսարանի», 2021, հ. 1 (58), էջ 229-237:
3. Սարգսյան Մ. Ա., Հայալեզու տեքստերի էլեկտրոնային նոր սրբագրիչի սկզբունքները, «Լեզու և լեզվաբանություն», հ. 2 (23), 2020, Երևան, էջ 10-17:
4. Սարգսյան Մ. Ա., Հայերեն էլեկտրոնային սրբագրման համակարգ, <https://armspell.am/hy>, 2021:

5. Advertisement (January 1981). "Micro Computer Industries, Ltd" (PDF). *Compute! Magazine, Issue 8, Vol. 3, No. 1.* p. 119.
6. Baeza-Yates R., Navarro G., A faster algorithm for approximate string matching. In Dan Hirschberg; Gene Myers. *Combinatorial Pattern Matching (CPM'96)*, LNCS 1075. Irvine, (June 1996). CA. pp. 1–23.
7. Banks, T. (2008). *Foreign Language Learning Difficulties and Teaching Strategies.* (pp. 29). Master's Thesis, Dominican University of California. Retrieved 19 March 2012.
8. Boytsov Leonid, Indexing methods for approximate dictionary searching: Comparative analysis. *Jea Acm*, 16 (1), 2011, pp. 1–91. doi: 10.1145/1963190.1963 191.
9. Bustamante F.R., León F.S, GramCheck: a grammar and style checker. *Proceedings of the 16th conference on Computational linguistics*, August 05–09, Copenhagen, Denmark, 1996, Retrieved 19 March 2012.

10. Dale, Robert; Moisl, Hermann; Somers, Harold (2000). Handbook of Natural Language Processing. CRC Press. ISBN 9780824790004.
11. de Amorim R.C., Zampieri M., Effective Spell Checking Methods Using Clustering Algorithms. Proceedings of Recent Advances in Natural Language Processing (RANLP2013). Hissar, Bulgaria, 2013, pp. 172-178.
12. Earnest, Les. "The First Three Spelling Checkers" (PDF). Stanford University. Archived from the original (PDF) on 22 October 2012. Retrieved 10 October 2011.
13. Earnest, Les. *Visible Legacies for Y3K* (PDF). Archived from the original (PDF) on 2011-07-20. Retrieved 2011-02-18.
14. *"Georgetown U Faculty & Staff: The Center for Language, Education & Development"*. Archived from the original on 2009-02-05. Retrieved 2008-12-18.
15. "Grammatik II". PC: The Independent Guide to IBM Personal Computers. Software Communications. 5: 190–199. 1986.

16. Harvey, Charlotte Bruce (May–June 2010). "Teaching Computers to Spell (obituary for Henry Kučera)". *Brown Alumni Magazine*. p. 79.
17. InfoWorld Media Group (28 October 1991). InfoWorld. InfoWorld Media Group, Inc. p. 68.
18. Kukich, K.: Techniques for Automatically Correcting Words in Text. R. Mitton, Ordering the suggestions of a spellchecker without using context, 2009.
19. Mitton, R.: Ordering the suggestions of a spellchecker without using context, Natural Language Engineering, 2009.
20. Peterson James,. Computer Programs for Detecting and Correcting Spelling Errors (PDF). (Dec 1980), Retrieved 2011-02-18.
21. Sargsyan M. A., Some Observations on Armenian Electronic Proofreading Systems, "World Science", №6 (34), Vol. 8, June 2018, Warsaw, Poland, pp. 21-24, DOI: https://doi.org/10.31435/rsglobal_ws/12062018/58

22. Smith, Charles R.; Kathleen E. Kiefer; Patricia S. Gingrich (1 July 1984). "Computers come of age in writing instruction". *Computers and the Humanities*. Springer Netherlands. 18 (3): 215–224. doi:10.1007/BF02267225. S2CID 28762117.
23. U.S. Patent 6618697, Method for rule-based correction of spelling and grammar errors, <https://patents.google.com/patent/US6618697>
24. <http://am.translit.cc/>
25. <http://am.translit.cc/>
26. http://www.cadastre.am/page/ashxarhagrakan_anvanumner
27. http://www.cadastre.am/page/ashxarhagrakan_anvanumner
28. <http://www.correctenglish.com/>
29. <http://www.gingersoftware.com/grammarcheck>
30. <http://www.hyspell.com/>
31. <http://www.hyspell.com/>
32. <http://www.onlinecorrection.com/>
33. <http://www.online-spellcheck.com/>

- 34. <http://www.powerspell.am/>
- 35. <http://www.powerspell.am/>
- 36. <http://www.reverso.net/spell-checker/english-spelling-grammar/>
- 37. http://www.stars21.com/spelling/armenian_spell_checker.html
- 38. http://www.stars21.com/spelling/armenian_spell_checker.html
- 39. <http://www.whitesmoke.com/>
- 40. <https://addons.mozilla.org/en-US/firefox/addon/armenian-spell-checker-diction/>
- 41. <https://addons.mozilla.org/en-US/firefox/addon/armenian-spell-checker-diction/>
- 42. <https://languagetool.org/>
- 43. <https://pro.spellcheckplus.com/>
- 44. <https://spellcheckplus.com/>
- 45. <https://www.grammarcheck.net/editor/>
- 46. <https://www.grammarly.com/>
- 47. <https://www.grammarly.com/>

48. <https://www.scribens.com/>
49. https://www.spellchecker.net/eastern_and_western_spell_checker.html
50. https://www.spellchecker.net/eastern_and_western_spell_checker.html

ՆՇՈՒՄՆԵՐԻ ՀԱՄԱՐ

ՆՇՈՒՄՆԵՐԻ ՀԱՄԱՐ

ՀՀ ԳԱԱ Հ. ԱՃԱՌՅԱՆԻ ԱՆՎԱՆ ԼԵԶԿԻ ԻՆՍՏԻՏՈՒՏ

ՄԵՐԻ ՍԱՐԳՍՅԱՆ

ՀԱՅԵՐԵՆ ԷԼԵԿՏՐՈՆԱՅԻՆ ՍՐԲԱԳՐՄԱՆ

ՀԱՄԱԿԱՐԳ

NAS RA INSTITUTE OF LANGUAGE AFTER H. ACHARYAN

MERI SARGSYAN

**ARMENIAN ELECTRONIC PROOFREADING
SYSTEM**

НАН РА ИНСТИТУТ ЯЗЫКА ИМ. Р. АЧАРЯНА

МЕРИ САРГСЯН

АРМЯНСКАЯ ЭЛЕКТРОННАЯ СИСТЕМА

КОРРЕКТУРЫ

Թուղթը՝ օֆսեթ:

Չափսը՝ 60X84/16

Ծավալը՝ 6 պայմ. տպ. մամուլ

Տպագրության եղանակը՝ թվային

Տպաքանակը՝ 150

Տպագրվել է «Տիր» հրատարակչության տպագրատանը
Հասցեն՝ Երևան, Ա Միկոյան փող., 2/2



9 789939 865911