

ՆՈՐ ԲԱՌԵՐԻ ՄԵՔԵՆԱԿԱՆ ՈՐՈՆՄԱՆ ԽՆԴԻՐՆԵՐԸ

ՖՐԻԴԱ ՀԱԿՈԲՅԱՆ

Վերջին տասնամյակներում հայ լեզվաբանության մեջ առավելապես կարևորվել է նորաբանությունների բառարանների ստեղծումը, ինչը պայմանավորված է հայերենի բառապաշարի արդի վիճակն արտացոլող բացատրական բառարանի բացակայությամբ: Այնուամենայնիվ, առկա նորաբանությունների բառարանները չեն արտացոլում նորաբանությունների ստեղծման ընդհանուր պատկերը, հիմնական օրինաչափությունները, քանի որ առանց տեխնիկական որևէ օժանդակության՝ անհնար է առանձնացել տեքստերում գործածված բոլոր նոր բառերը: Իբրև խնդրի լուծում՝ առաջարկվում է ստեղծել բառապաշարի համակարգչային շտեմարան և համացանցային տարբեր աղբյուրներից մեքենական եղանակով համալրվող տեքստային հենք:

Բանալի բառեր – բառապաշարի համակարգչային շտեմարան, մոնիտորային կորպուս, բառարան, նորաբանություն, նորաբանությունների բառարան, տեքստային հենք:

Արդի հայ բառարանագրության առջև դրված խնդիրներից հարկ է էապես կարևորել ինչպես հայերենի բառապաշարի համաժամանակյա և տարաժամանակյա ուսումնասիրությունը, բառերի իմաստային և հասկացական կողմերի բացահայտումն ու քննությունը, այնպես էլ լեզվի բառապաշարի համակարգչային շտեմարաններ, տվյալների հենքեր ստեղծելու և դրանք նորաբանություններով, նոր բառերով մեքենական եղանակով հարստացնելու հարցերը:

Ինչպես հայտնի է, նորաբանությունները բառային այն միավորներն են, այն նոր բառերը, արտահայտությունները և իմաստները, որոնք լեզվի տվյալ փուլում գիտակցվում են իբրև նոր իրողություններ, դեռևս չեն մտել գործուն բառապաշարի մեջ, չունեն ընդհանուր տարածում, լայն և հաճախական կիրառություն, իսկ առանձին դեպքերում նույնիսկ չեն արձանագրվել տվյալ լեզվի համապատասխան բառարաններում¹:

Հատկապես վերջին տասնամյակներում հայ լեզվաբանության մեջ առանձնահատուկ ուշադրություն է դարձվել նորաբանությունների և նոր բառերի տարբերակման խնդրին, դրանց իմաստային քննությանը, բառակազմական կաղապարների վերլուծությանը: Առավելապես կարևորվել է նաև նորաբանությունների բառարանների ստեղծումը, ինչը պայմանավորված է հայե-

¹ Տե՛ս **Եզեկյան Լ.**, Հայոց լեզվի ոճագիտություն, Երևան, 2007, էջ 132:

րենի բառապաշարի արդի վիճակն արտացոլող բացատրական բառարանի բացակայությամբ, քանի որ համեմատաբար վերջին շրջանում ստեղծված բացատրական բառարաններից ԳԱԱ Հ. Աճառյանի անվան լեզվի ինստիտուտի կազմած «Ժամանակակից հայերենի բացատրական բառարանի» (1969-1980), Է. Աղայանի «Արդի հայերենի բացատրական բառարանի» (1976), Հ. Բարսեղյանի «Հայերեն ուղղագրական, ուղղախոսական, տերմինաբանական բառարանի» (1973) լույսընծայումից անցել է մի քանի տասնամյակ: Այդ ընթացքում լեզվի բառապաշարը կրել է նկատելի փոփոխություններ. մի կողմից՝ դուրս են մնացել բազմաթիվ հնաբանություններ, սակավ գործածական բառեր, մյուս կողմից՝ բառապաշարը համալրվել է հսկայական թվով նորաբանություններով, որոնցից շատերը տարիների ընթացքում լայնորեն գործածվել են և համալրել հայերենի բառապաշարի հարուստ բառազանձը:

Վերոնշյալ բացատրական բառարանները չեն ներառում լեզվում հաճախադեպ կիրառություն ունեցող բոլոր նորակազմությունները, ուստի ժամանակ առ ժամանակ հրատարակվում են նորաբանությունների բառարաններ: Մինչև օրս ստեղծված նորաբանությունների բառարաններից կարող ենք առանձնացնել Ֆ. Խլղաթյանի «Բառարան-տեղեկատուն (նոր բառեր և իմաստներ)» (1982), որն ընդգրկում է 70-80-ական թթ. ժամանակակից հայերենում ստեղծված (մասամբ նաև նախորդ բառարաններից դուրս մնացած) և այդ տարիների պարբերական մամուլում գործածված բառերը: Բառարանից դուրս են մնացել նորագույն գեղարվեստական գրականության մեջ գործածված, սակայն լայն կիրառություն չունեցող հեղինակային նորակազմությունները. լեզվաբանի գլխավոր նպատակը եղել է հաճախակի գործածվող բառերը ներկայացնելը: Հաջորդը Փ. Մեյթիխանյանի «Նոր բառերի բացատրական բառարան» է, որը պարունակում է վերջին տարիներին մամուլում և գեղարվեստական գրականության մեջ գործածված շուրջ 3000 նոր բառեր ու տերմիններ²:

Նորաբանությունների բառարանների թվում առանձնանում է Ս. Էլոյանի «Արդի հայերենի նորաբանությունների բառարանը» (2002): Ունենալով շուրջ 30 տարվա ժամանակահատվածում հայ մամուլի և գեղարվեստական գրականության մեջ գործածված նորաբանությունների հարուստ բառաքանակ՝ լեզվաբանն իր բառարանում ընդգրկել է բառագիտական տեսակետից բոլոր տիպի նորաբանությունները՝ նորակազմ բառերը, տերմինները (ինչպես փոխառյալ, այնպես էլ սեփական բառակազմական կաղապարներով կազմված),

² Էլոյան Ս., Արդի հայերենի նորաբանությունների բառարան, Երևան, 2002, էջ 7:

իմաստային նորաբանությունները (արդեն եղած բառերի նոր ստացած իմաստները), նոր դարձվածքները, կայուն և ազատ բառակապակցությունները, հապավումները (ոչ զուտ տառային), առանձնահատուկ արտահայտությունները, ինչպես նաև անհատական-ոճական նորաբանություններն իրենց դիպվածական և հեղինակային տեսակներով³:

Վերոնշյալ բառարանների համեմատությամբ իրենց ընդգրկած բառապաշարի ծավալով և բազմազանությամբ, պարբերաբար հրատարակվելու շնորհիվ առավել մեծաթիվ բառաքանակ ներկայացնելու հնարավորությամբ շահեկանորեն տարբերվում են ԳԱԱ Հ. Աճառյանի լեզվի ինստիտուտում պարբերաբար հրատարակվող «Նոր բառերի» պրակները: Պրակներում ներկայացված են համացանցից, մամուլից, գիտական և գեղարվեստական գրականությունից, հետուստատեսությունից, գովազդային հոլովակներից ու ցուցանակներից քաղված մեծաթիվ բառեր, որոնք վկայված չեն եղած բացատրական և նորաբանությունների բառարաններում: Պրակների հեղինակները նպատակ ունեն հավաքված հազարավոր նորակազմություններից, որոշակի սկզբունքով բառընտրություն կատարելով, կազմել «Նոր բառերի ընդարձակ բառարան», որում կարտացոլվեն XXI դարասկզբի հայերենի բառապաշարում տեղի ունեցած տեղաշարժերը⁴:

Մեծապես կարևորելով այս և հետագա շրջանում նորաբանությունների ու նոր բառերի համակողմանի քննությունն ու համապատասխան բառարանների ստեղծումը, կարծում ենք՝ արդի հայ բառարանագրության առաջնային խնդիրը ոչ թե նորաբանությունների, այլ հայերենի նորագույն բացատրական բառարանի ստեղծումն է: Տրամաբանական չէ քառորդ դար առաջ ստեղծված և լայն գործածություն ունեցող բառերը նորաբանություն համարել միայն այն պատճառով, որ դրանք վկայված չեն կես դար առաջ ստեղծված բացատրական բառարաններում: Նորագույն բացատրական բառարանի նպատակը պետք է լինի բառապաշարի համաժամանակյա վիճակի արտացոլումը, հետևաբար հայերենի հիմնական բառաֆոնդին պատկանող բառերին զուգահեռ վերջին տասնամյակներում ստեղծված հազարավոր բառերը կներկայացվեն ոչ թե որպես նորաբանություն, այլ հայերենի բառապաշարի մաս: Համակարգչային բառարանագրության զարգացման այս շրջանում միայն

³ Տե՛ս նույն տեղում, էջ 6:

⁴ Գալստյան Ա., Հովսեփյան Գ., Սահակյան Ա., Սահինյան Լ., Նոր բառեր. Ա պրակ, Երևան, 2015, էջ 4:

Նորագույն բացատրական տպագիր բառարանն այլևս բավական չէ. ինչպես հայտնի է, համակարգչային տեխնիկան լեզվաբան-բառարանագիրներին առաջարկում է բառապաշարի մշակման և ներկայացման առավել լայն հնարավորություններ: Ուստի պետք է կարևորել բացատրական բառարանի հիման վրա հայերենի բառապաշարի լիակատար համակարգչային շտեմարան ստեղծելու հանգամանքը: Չխոսելով համակարգչային շտեմարանի՝ տպագիր բառարանի համեմատությամբ ունեցած նկատելի տարբերությունների մասին, նշենք միայն, որ այն հնարավոր կլինի շարունակ հարստացնել-համալրել նոր բառերով ու նորաբանություններով՝ առանց ժամանակային որևէ սահմանափակման: Ուստի հարկ կլինի նաև վերանայել նորաբանությունների ավանդական ըմբռնման հարցը, քանի որ նորաբանություն են համարվում բառարաններում դեռևս չընդգրկված բառերը, իսկ շտեմարանում կարելի է դրանք ավելացնել ցանկացած պահի: Ակնհայտ է, որ համակարգչային շտեմարանը կարող է լիովին արտացոլել բառապաշարի արդի վիճակը ցանկացած ժամանակահատվածի հարաբերությամբ՝ ի տարբերություն տպագիր բառարանի, որի հրատարակումից անցնում են տասնամյակներ, և առաջանում է զգալի անջրպետ տվյալ փուլում լեզվի բառապաշարի և տպագիր բառարանում ներկայացված բառապաշարի միջև: Ավելին՝ շտեմարանն առավել լիակատար կլինի, եթե ներկայացնի ոչ միայն գրական արևելահայերենի, այլև արևմտահայերենի բառապաշարը: Պետք է մեծապես կարևորել նաև արևելահայերեն-արևմտահայերեն առցանց բառարան ստեղծելու հանգամանքը՝ գրական հայերենի այս երկու ճյուղերի մերձեցման ուղիների որոնման համատեքստում:

Պակաս կարևոր չէ նաև նորաբանությունների բառարանները հարստացնելու, տեքստերում գործածված նորակազմությունների որոնման գործընթացն առավելագույն չափով մեքենայացնելու խնդիրը:

Հարկ է նկատել, որ վերջին տասնամյակներում ստեղծված նոր բառերի և նորաբանությունների իրական թիվն անհամեմատ ավելի մեծ է, եթե այն համեմատենք նորաբանությունների բոլոր բառարաններում ընդգրկված բառաքանակի հետ: Ուստի բովանդակային առումով առկա նորաբանությունների բառարանների հիմնական թերությունն այն է, որ դրանցում ներկայացված են լեզվում ստեղծված մեծաթիվ նորաբանությունների միայն փոքր մասը: Ոչ համաչափ են ընտրված նաև այն տպագիր և էլեկտրոնային աղբյուրները, որոնցից քաղվում են նորաբանությունները:

Նորաբանությունների բառարանները չեն արտացոլում նորաբանությունների ստեղծման ընդհանուր պատկերը, հիմնական օրինաչափություն-

ները, քանի որ առանց տեխնիկական որևէ օժանդակության՝ անհնար է առանձնացել տեքստերում գործածված բոլոր նորաբանությունները: Հետևաբար՝ հատկապես հիմա լեզվաբանների կողմից նորաբանությունների որոնման, քարտագրման ավանդական եղանակն այլևս չի կարող համարվել նորաբանությունների բառարանների կազմման հուսալի ճանապարհ:

Եթե նախկինում տեքստի հանրային տարածման միակ եղանակը դրա տպագրությունն էր, այժմ համացանցի ընձեռած լայն հնարավորությունների շնորհիվ նկատելիորեն հեշտացել է էլեկտրոնային տեքստերի կազմումը և դրանց հրապարակումը համացանցում: Ասվածի լավագույն ապացույցը հարյուրավոր հայալեզու կայքերի առկայությունն է: Նույնիսկ վերջին շրջանում ստեղծված նորաբանությունների բառարանների հիմքում առավելապես տպագիր աղբյուրներն են, ընդ որում՝ նորաբանությունները քաղված են հիմնականում գեղարվեստական գրականությունից, գրականագիտական, լեզվաբանական ուսումնասիրություններից, ամսագրերից («Անդին», «Աստղիկ», «Արտասահմանյան գրականություն», «Գարուն», «Լրաբեր հասարակական գիտությունների», «Նորք», «Պատմաբանասիրական հանդես») և լրագրերից («Ազրուխտու», «Ազգ», «Գրական թերթ», «Երևանի համալսարան» ամսաթերթ, «Իրատես», «Հայաստանի Հանրապետություն», «Հրապարակ»)⁵: Մինչդեռ համացանցային ամենահայտի և տարածված լրատվական ու մեդիա կայքերը բացակայում են բառարանների օգտագործված գրականության ցանկերից: Կարծում ենք՝ նորաբանությունների որոնումը պետք է սկսել հենց այդ կայքերի էլեկտրոնային տեքստերից, քանի որ դրանցում նույնիսկ օրվա կտրվածքով կարելի է գտնել լայն և հաճախակի կիրառություն ունեցող մեծաթիվ նորաբանություններ՝ ի տարբերություն գեղարվեստական գրականության մեջ սակավ գործածություն ունեցող հեղինակային նորակազմությունների:

Հասարակական կյանքի գրեթե բոլոր բնագավառներում տեղի են ունենում արմատական վերափոխումներ, որոնք էլ իրենց հերթին հանգեցրել են մեծ թվով նորաբանությունների ստեղծմանը: Այնուամենայնիվ, դրանցից նույնիսկ հաճախադեպ գործածվողները բացակայում են նորաբանությունների բառարաններից: Օրինակ՝ *վերջին շրջանում էլեկտրոնային տարբեր համակարգերում լայնորեն օգտագործվում են այնպիսի վճարամիջոցներ, ինչպիսիք են կրիպտոարժույթները: 2013 թվականից բիթքոյինը դարձավ ամբողջ աշխարհում շրջանառվող ամենատարածված կրիպտոարժույթը: Կան կրիպտո-*

⁵ Տե՛ս նույն տեղում, էջ 151-155:

ծառայություններ մատուցող բազմաթիվ ընկերություններ: **Կրիպտոարժույթ** ստանալու ողջ գործընթացը կոչվում է **մայնինգ**, այն մասնագիտացված համակարգիչները, որոնք ապահովում են **բիթքոյինների** ցանցի անվտանգությունը, հաստատում են վճարումները և թողարկում թվային նոր դրամներ, անվանվում են **մայներ**: **Բիթքոյինով** կատարվող գործարքները գրանցելու եւ հաստատելու համար գործածվող թվային տեխնոլոգիան կոչվում է **բլոկչեյն** (բլոկների շղթա): **Մայներ**, **մայնինգ**, **բլոկչեյն** օտարաբանությունները հայերեն համարժեքներով փոխարինելու խնդիր կա, իսկ **կրիպտոարժույթ**, **կրիպտովալյուրա**, **կրիպտոարժույթային**, **կրիպտոծառայություն**, **կրիպտոանվտանգություն**, **կրիպտովիրուս** բառերը, որոնք էլեկտրոնային մամուլում ունեն լայն և հաճախադեպ կիրառություն, նորաբանությունների բառարաններում ընդգրկված չեն: Վերոնշյալ մասնավոր օրինակներն ընդգծում են էլեկտրոնային տեքստերից նորաբանությունների և նոր բառերի որոնումն առավելագույն չափով լիակատար դարձնելու և մեքենայացնելու կարևորությունը: Դրա միակ եղանակը հայերենի մոնիտորային կորպուսի ստեղծումն է: Ընդհանուր առմամբ՝ կորպուսները հսկայական նշանակություն ունեն լեզվական փաստերի քննության համար, քանի որ պարունակում են տասնյակ հազարավոր տեքստեր: Այս դեպքում մոնիտորային կորպուսը ոչ միայն հնարավոր կդարձնի կատարել արդի հայերենի լեզվական բոլոր իրողությունների համակողմանի ուսումնասիրություն տեքստային մեծաթիվ օրինակներով, այլև էապես կդյուրացնի բառարանագրական աշխատանքները: Համացանցային զանազան աղբյուրներից մեքենական եղանակով հարստացվող տեքստային հենքի առկայությունը հնարավոր կդարձնի հստակ տեսնել արդի հայերենի բառապաշարում տեղի ունեցող փոփոխությունները, քննել նորաստեղծ բառերը, դրանց իմաստային փոփոխությունները, ճշտել գործածության ոլորտները: Ավելորդ է նշել, թե որքան կմեծանա մոնիտորային կորպուսի կարևորությունը, եթե այն դառնա զուգադրական կորպուս՝ էլեկտրոնային աղբյուրներից համարվելով ոչ միայն արևելահայերեն, այլև արևմտահայերեն տեքստերով:

Ինչ խոսք, արդի փուլում հայ բառարանագրության առջև դրված խնդիրները չեն սահմանափակվում միայն նոր բառերի մեքենական որոնմամբ և համապատասխան բառարանների ստեղծումով: Առավել լուրջ և հեռանկարային ենք համարում հայերենի պատմական էլեկտրոնային բառարանի ստեղծումը, որը կներկայացնի ոչ միայն նոր բառերը, այլև հայերենի ողջ բառապաշարը լեզվի պատմական զարգացման բոլոր փուլերում՝ մեծաթիվ և բազմազան աղբյուրներից քաղված օրինակներով: Այդ նպատակի իրականացմանը լիովին

կօժանդակի այս դեպքում արդեն ոչ թե մոնիտորային կամ զուգադրական, այլ հայերենի ընդհանրական կորպուսը: Նշենք, որ ընդհանրական կորպուսի ստեղծման փորձ արվել է. մեր կողմից ստեղծված, hamabarbar.am կայքում գործող էլեկտրոնային համաբարբառների մեծածավալ շտեմարանը կարող է համարվել ընդհանրական **կորպուս-համաբարբառ**, քանի որ այն ներկայացնում է գրաբարալեզու և ժամանակակից բոլոր խոշորագույն հեղինակների երկերի բառապաշարը բնագրային գործածության համապատասխան վկայություններով՝ տալով բառերի, բառաձևերի, արտահայտությունների և դարձվածքների որոնման հնարավորություն: Հետևաբար առկա նորաբանությունների բառարանների հարստացման մեկ այլ աղբյուր կարող է համարվել հայ նոր և նորագույն գրականության ներկայացուցիչների երկերում գործածված նորակազմություններով համալրումը:

Այսպիսով՝ բառապաշարի համակարգչային շտեմարանի և համացանցային տարբեր աղբյուրներից մեքենական եղանակով ինքնահամալրվող տեքստային հենքի ստեղծումը արմատապես կփոխի բառարանագրական աշխատանքների բնույթը, ընդգրկման ծավալը և զարգացման հեռանկարները:

ПРОБЛЕМЫ АВТОМАТИЧЕСКОГО ПОИСКА НОВЫХ СЛОВ

ФРИДА АКОПЯН

В последние десятилетия лингвисты уделяют особое внимание созданию словарей неологизмов. Текущие словари неологизмов не отражают общую картину неологизмов, потому что без какой-либо технической помощи невозможно получить все неологизмы, используемые в текстах. В качестве решения проблемы рекомендуется создать лексическую базу данных и корпус, обогащенный различными веб-источниками.

Ключевые слова – база данных, мониторинг корпус, словарь, неологизм, словарь неологизмов.

THE PROBLEMS OF THE AUTOMATIC SEARCH OF NEW WORDS

FRIDA HAKOBYAN

In the last decades linguists have paid a special attention to the creation of dictionaries of neologisms. It is conditioned by the lack of a dictionary reflecting the current state of the Armenian vocabulary. Nevertheless, the current

dictionaries of neologisms do not reflect the general picture of neologisms because without any technical assistance it is impossible to get all neologisms used in texts. As a solution to the problem, it is suggested to create a lexical database and a monitor corpus enriched with a variety of web sources.

Key words – lexical database, dictionary, neologisms, dictionary of neologisms, corpus, monitor corpus.