

**ՀԱՅԱՍՏԱՆԻ ՀԱՆՐԱՊԵՏՈՒԹՅԱՆ
ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԶԳԱՅԻՆ ԱԿԱԴԵՄԻԱ
ՀՐԱԶՅԱ ԱՃԱՌՑԱՆԻ ԱՆՎԱՆ ԼԵԶՎԻ ԻՆՍՏԻՏՈՒՏ**

ՖՐԻԴԱ ԱՐՏԱՇԵՄԻ ՀԱԿՈՔՑԱՆ

**ԲԱՌԱՄԻԱՎՈՐՆԵՐԻ ՆԿԱՐԱԳՐՈՒԹՅԱՆ ՄԿԶԲՈՒՆՔՆԵՐԸ ԺԱՄԱՆԱԿԱԿԻՑ
ՀԱՅՈՑ ԼԵԶՎԻ ԲԱՌԱՊԱՇԱՐԻ ՀԱՄԱԿԱՐԳՉԱՅԻՆ ՇՏԵՄԱՐԱՆՈՒՄ**

Ժ.02.02 - <<Ընդհանուր և համեմատական լեզվաբանություն>> մասնագիտությամբ բանասիրական գիտությունների թեկնածուի գիտական աստիճանի հայցման ատենախոսության

ՍԵՂՄԱԳԻՐ

ԵՐԵՎԱՆ 2015

Ատենախոսության թեման հաստատվել է ՀՀ ԳԱԱ Հրաչյա Աճառյանի անվան լեզվի ինստիտուտում

Գիտական ղեկավար՝

բանասիրական գիտությունների դոկտոր, պրոֆեսոր

Լ. Ս. Հովսեփյան

Պաշտոնական ընդդիմախոսներ՝

բանասիրական գիտությունների դոկտոր, պրոֆեսոր

Ս.Ա. Հայրապետյան

բանասիրական գիտությունների թեկնածու, դոցենտ

Ա. Լ. Մարգարյան

Առաջատար կազմակերպություն՝

Երևանի Վ. Բրյուսովի անվան պետական լեզվահասարակագիտական համալսարան

Պաշտպանությունը կայանալու է 2016 թվականի հունվարի 25-ին, ժամը՝ 15:00-ին, ՀՀ ԳԱԱ Հ. Աճառյանի անվան լեզվի ինստիտուտում գործող ՀՀ ԲՈՀ-ի՝ Լեզվաբանության 019 մասնագիտական խորհրդում

Հասցեն՝ 0015, ք. Երևան, Գրիգոր Լուսավորչի 15:

Ատենախոսությանը կարելի է ծանոթանալ ՀՀ ԳԱԱ Հ. Աճառյանի անվան լեզվի ինստիտուտի գրադարանում:

Սեղմագիրն առաքված է 2015թ. դեկտեմբերի 21-ին:

Մասնագիտական խորհրդի գիտական քարտուղար, բանասիրական գիտությունների թեկնածու, դոցենտ

Ն.Մ.Միսնոյան

ԱՏԵՆԱԽՈՍՈՒԹՅԱՆ ԸՆԴՀԱՆՈՒՐ ԲՆՈՒԹԱԳԻՐԸ

ԹԵՄԱՅԻ ԱՐԴԻԱԿԱՆՈՒԹՅՈՒՆԸ: Ատենախոսության թեմայի արդիականությունը պայմանավորված է առավել լիակատար կերպով հայերենը տեղեկատվական տեխնոլոգիաների բնագավառում կիրառելու անհրաժեշտությամբ, որը կծառայի երկու նպատակի: Մի կողմից՝ համակարգչային տեխնիկական միջոցների նորագույն նվաճումներն օգտագործելով՝ հնարավոր կլինի նոր ձևերով կատարել հայերենի լիակատար ուսումնասիրությունը, համակարգել լեզվական նյութը, լեզվական իրողությունները հետազոտել և համապատասխան ձևով՝ թվային տարբեր միջոցներով ներկայացնել, մյուս կողմից՝ համացանցում եղած հսկայածավալ և տարաբնույթ տեղեկատվությունը հայերենով հասանելի դարձնել հայ օգտվողներին: Քանի որ ցանակացած լեզվի հիմնական առաջը այդ լեզվի բառապաշարն է, լեզուն համակարգչի միջոցով ուսումնասիրելն ու ներկայացնելը ամենից առաջ ենթադրում են այդ լեզվի բառապաշարի ներկայացում: Հետևաբար հայերենը համացանցային լեզու դարձնելու և այն տեղեկատվական տեխնոլոգիաների բնագավառում առավել հաջողությամբ գործածելու համար առաջին հերթին անհրաժեշտ է նկարագրել հայերենի բառապաշարը՝ օգտագործելով ժամանակակից համակարգչային տեխնիկայի անընդհատ ու շարունակական զարգացումները: Այդ նպատակին հասնելու լավագույն եղանակը հայերենի բառապաշարի համակարգչային շտեմարան կազմելն է:

ԱՏԵՆԱԽՈՍՈՒԹՅԱՆ ՆՊՏԱԿՆ ՈՒ ԽՐԴԻՐՆԵՐԸ: Ատենախոսության նպատակն է, մի կողմից՝ հենվելով ժամանակակից համակարգչային բառարանագրության ուսանելի փորձի վրա, մյուս կողմից՝ հաշվի առնելով հայերենի լեզվական առանձնահատկությունները, մշակել հայերենի բառապաշարի համակարգչային շտեմարան կազմելու սկզբունքները, որոնց համաձայն ներկայացված կլինի հայերենի բառապաշարն իր համաժամանակյա վիճակով, կկատարվի բառապաշարի հնարավորինս ամբողջական իմաստային դասակարգում, ներկայացված կլինեն բառերի հասկացական և իմաստային փոխհարաբերությունները, կտրվի բառի ձևաիմաստային սպառիչ բնութագիրը:

ԱՏԵՆԱԽՈՍՈՒԹՅԱՆ ՏԵՍԱԿԱՆ-ՄԵԹՈԴԱԲԱՆԱԿԱՆ ՀԻՄՔԸ: Ատենախոսության մեթոդաբանական հիմքում ընկած են ժամանակակից լեզվաբանության փորձն ու արդի գիտության նվաճումները: Կիրառվել են կառուցվածքային լեզվաբանության համաժամանակյա ուսումնասիրության մեթոդները, համակարգչային լեզվաբանության մեջ լեզվական հետազոտության ներքին սկզբունքներն ու լեզվի ձևային նկարագրության եղանակները:

ԱՏԵՆԱԽՈՍՈՒԹՅԱՆ ԳԻՏԱԿԱՆ ՆՈՐՈՒՅԹՆ ՈՒ ԳՈՐԾՆԱԿԱՆ ՆՇԱՆԱԿՈՒԹՅՈՒՆԸ: Հաշվի առնելով այն հանգամանքը, որ ներկայումս համակարգչային բառարանագիրներին դեռևս չի հաջողվել ո՛չ էլեկտրոնային բառարանների, ո՛չ բառային հենքերի միջոցով ներկայացնել բառի լիարժեք ձևաիմաստային բնութագիրը, ատենա-

խոսության մեջ առաջին անգամ նախ՝ առաջարկվել են հայերենի բառապաշարի համակարգչային շտեմարանի կազմության սկզբունքները, ապա՝ փորձ է արվել դրանց հիման վրա նոր կազմվելիք շտեմարանում հնարավորինս սպառիչ ու ամբողջական կերպով տալ գլխաբառի հնչյունական, իմաստային ու քերականական բնութագիրը: Շտեմարանը կծառայի ինչպես տեսական-լեզվաբանական հարցերի ուսումնասիրությանը, այնպես էլ գործնական-կիրառական խնդիրների լուծմանը: Այն կծառայի թե՛ մասնագետ-լեզվաբաններին, որոնք շտեմարանից կստանան լեզվաբանական տարաբնույթ օգտակար տեղեկատվություն հետազոտական աշխատանքներ կատարելիս, թե՛ հայերեն իմացող յուրաքանչյուր մարդու, որը կուզենա շտեմարանից ցանկացած բառի մասին բառային տեղեկատվություն ստանալ, թե՛ այն մարդկանց, որոնք հայերենին լավ չեն տիրապետում և ցանկանում են հարստացնել իրենց լեզվական գիտելիքները: Կարող ենք վստահորեն ասել, որ լեզվաբանական նպատակներով շտեմարանի կիրառությունը զարգացման միանգամայն նոր աստիճանի ու որակի կհասցնի հայերենի գործածությունը տեղեկատվական տեխնոլոգիաների բնագավառում: Առաջարկվող սկզբունքներով կազմվելիք շտեմարանը կօժանդակի հետազոտական բազմազան հարցերի ուսումնասիրությանը, կլինի ավելի լիակատար և առավել հաջողությամբ կծառայի ինչպես մեքենական թարգմանության, այնպես էլ գործնական-կիրառական ծրագրերի իրագործմանը (տեղեկատվության ավտոմատ մշակում, ավտոմատ սրբագրություն և խմբագրություն և այլն):

ԱՏԵՆԱԽՈՍՈՒԹՅԱՆ ՓՈՂՉԱՔՆՆՈՒԹՅՈՒՆԸ: Ատենախոսությունը քննարկվել և հրապարակային պաշտպանության է երաշխավորվել ՀՀ ԳԱԱ Հ. Աճառյանի անվան լեզվի ինստիտուտի ընդհանուր և համեմատական լեզվաբանության բաժնի կողմից: Թեմայի հիմնադրույթներն արտացոլված են գիտական տարբեր հանդեսներում հեղինակի տպագրած չորս հոդվածներում:

ԱՏԵՆԱԽՈՍՈՒԹՅԱՆ ԿԱՌՈՒՅՎԱԾՔՆ ՈՒ ԲՈՎԱՆԴԱԿՈՒԹՅՈՒՆԸ: Ատենախոսության կառուցվածքը բխում է առաջադրված նպատակներից, քննարկվող խնդիրներից և արդյունքներից: Այն բաղկացած է ներածությունից, երեք գլխից, եզրակացություններից և օգտագործված գրականության ցանկից: Ատենախոսության ծավալն է 144 էջ:

Ներածության մեջ ընդհանուր գծերով ներկայացրել ենք համակարգչային լեզվաբանությունն իր ամենահիմնական բնութագրական հատկանիշներով, անդրադարձել համակարգչային լեզվաբանության տեսական ու կիրառական խնդիրներին, համակարգչի միջոցով լեզուն ուսումնասիրել-ներկայացնելու համար անհրաժեշտ լեզվի ձևային նկարագրությանը, լեզվի բառապաշարի ներկայացման եղանակներին, ինչպես նաև աշխատանքի նպատակին ու խնդիրներին, թեմայի արդիականությանն ու հարակից այլ հարցերի:

Ատենախոսության առաջին՝ **«Ժամանակակից հայոց լեզվի բառապաշարի համակարգչային շտեմարանի կազմման նախադրյալները»**, գլխում քննել ենք համացանցում բառային տեղեկատվության ներկայացման եղած աղբյուրները՝ պարզելով, թե

արդյո՞ք դրանք կարող են համակարգչային շտեմարան համարվել և որպես օրինակ ծառայել հայերենի համակարգչային շտեմարանի կազմության գործում:

Ժամանակակից համակարգչային տեխնիկայի ընձեռած լայն հնարավորությունները զարգացման միանգամայն նոր աստիճանի հասցրին համակարգչային բառարանագրության շրջանակներում կատարվող ամենատարբեր բնույթի աշխատանքները: Համակարգչային բառարանագրությունն աստիճանաբար դառնում է համակարգչային լեզվաբանության ամենաարագ զարգացող ուղղություններից մեկը: Այն զբաղվում է էլեկտրոնային բառարանների ստեղծման սկզբունքների, բառահոդվածների կազմման ու ներկայացման եղանակների մշակմամբ:

Քանի որ համացանցում հասանելի էլեկտրոնային բառարաններն ունեն կազմության ու բառամիավորների ներկայացման սկզբունքների ընդհանրություններ, դրանցից յուրաքանչյուրին առանձին-առանձին անդրադառնալու փոխարեն ներկայացրել ենք մի քանի տասնյակի հասնող այդ բառարանների միջև եղած ընդհանրություններն ու տարբերությունները, առանձնացրել իրենց կազմության սկզբունքներով առավել օրինակելիները, խոսել տպագիր և էլեկտրոնային բառարանների նմանությունների ու տարբերությունների մասին, նշել այն նպատակները, որոնց ծառայում են էլեկտրոնային բառարանները ուրվագծելով արտասահմանյան էլեկտրոնային բառարանագրության զարգացման արդի միտումները: Այնուհետև ներկայացրել ենք առավել գործածական միալեզվյան ու բազմալեզվյան բառարանները՝ ստեղծված ինչպես անգլերենի, այնպես էլ աշխարհի մի շարք այլ լեզուների համար: Անգլերենի համար ստեղծված էլեկտրոնային բառարաններից առանձնացրել ենք www.collinsdictionary.com կայքում ներկայացված Collins Unabridged English Dictionary, Collins Webster's American English Dictionary¹ բառարանները, www.ldoceonline.com կայքում ներկայացված Longman Dictionary of Contemporary English բառարանը², www.merriam-webster.com կայքում ներկայացված Merriam-Webster's Collegiate Dictionary բառարանը³, www.oxforddictionaries.com կայքում ներկայացված կայքում ներկայացված Օքսֆորդի՝ բրիտանական և ամերիկյան անգլերենների բառարանները, www.macmillandictionary.com կայքում ներկայացված Macmillan English Dictionary for Advanced Learners բառարանները⁴:

Ներկայացրել ենք ինչպես անգլերենի, այնպես էլ մի շարք այլ լեզուների համար ստեղծված առցանց բառարանները, որոնցից ամենատարածվածներն են լիտվերենի **Academic Dictionary of Lithuanian** (Didysis lietuvių kalbos žodynas կամ Akademinis lietuvių kalbos žodynas) բառարանը, հոլանդերենի **Woordenboek der Nederlandsche Taal** և **Van Dale Groot woordenboek van de Nederlandse taal** բառարանները, շվեդերենի **Svenska Akademiens ordbok** (SAOB) բառարանը, որը հրատարակվել է Շվեդիայի

¹ Տե՛ս <http://www.collinsdictionary.com>:

² Տե՛ս <http://www.ldoceonline.com>:

³ Տե՛ս <http://www.merriam-webster.com>:

⁴ Տե՛ս <http://www.macmillandictionary.com>:

գիտությունների ակադեմիայի կողմից Ordbok över svenska språket utgiven av Svenska Akademien վերնագրով⁵:

Ընդհանուր առմամբ գնահատելով օտարալեզու էլեկտրոնային բառարանները և քննելով դրանց օրինակելի լինելը հայ համակարգչային բառարանագիրների համար՝ նկատել ենք, որ այսօր գոյություն ունեցող էլեկտրոնային ոչ բոլոր տեսակի բառարաններն են կազմված գիտական բարձր չափանիշներով: Նախ օտարալեզու բառարաններում ոչ միշտ են ամբողջական կերպով ներկայացվում բառամիավորները, հետո՝ չենք կարող նաև լիովին վստահ լինել, թե դրանք կազմված են բացառապես լեզվաբան-բառարանագիրների կողմից: Նկատենք միայն, որ հաճախ այդ բառարանները համարվում են համակարգչային զանազան ծրագրերի միջոցով տեքստից քաղված բառերի ինքնահարստացման եղանակով: Բացի այդ դրանցում նոր բառեր կարող են ավելացնել նաև սովորական օգտատերերը: Այդ դեպքում դժվար է վստահ լինել նրանց լեզվական գիտելիքների վրա, մանավանդ եթե հաշվի առնենք, թե որքան դժվար է բառարանի բառահողոված կազմելը նույնիսկ ամենահմուտ բառարանագրի համար: Այդ ամենը հաշվի առնելով էլ՝ կարևորել ենք, որ հայ համակարգչային բառարանագիրները, օգտվելով էլեկտրոնային բառարանագրության արտասահմանյան փորձից, պետք է ոչ թե մեխանիկորեն կրկնօրինակեն, այլ մշակեն նոր սկզբունքներ՝ հայերեն էլեկտրոնային բառարաններ կազմելու համար:

Արտասահմանյան էլեկտրոնային բառարանները ներկայացնելուց հետո անդրադարձել ենք հայ համակարգչային բառարանագրության ոլորտում կատարվող աշխատանքներին ընդհանրապես (հաճախ արվում են ոչ արհեստավարժ լեզվաբան-բառարանագիրների կողմից և մասնավոր բնույթ են կրում), և հայերեն էլեկտրոնային բառարաններին՝ մասնավորապես:

Ավելին՝ դժվար է նույնիսկ վստահորեն ասել, թե մենք ունենք հայերենի էլեկտրոնային բառարաններ: Այսպես՝ այսօր համացանցային տարբեր կայքերում հասանելի այսպես կոչված հայերեն էլեկտրոնային բառարանները կամ տպագիր բառարանների թվային տարբերակներն են, կամ այսպես կոչված նորակազմ բառարաններ (մեծ մասամբ թարգմանական), որոնցում տրվում է բառամիավորի կարճ և ընդհանուր բացատրությունը: Որչափով սրանք կարող են լիակատար էլեկտրոնային բառարան համարվել՝ դժվար է ասել: Ճիշտ է, սկզբնական շրջանում էլեկտրոնային էին համարվում տպագիր բառարանների թվայնացված տարբերակները, և եթե մենք էլ այդ թվային տարբերակներն էլեկտրոնային բառարան համարենք, ապա կարող ենք ընդունել, որ դրանք էլեկտրոնային բառարաններ են: Մինչդեռ հիմա համակարգչային բառարանագրության մեջ էլեկտրոնային բառարան ասելով՝ հասկանում ենք նոր սկզբունքներով ու եղանակներով կազմված, համակարգչային զանազան ծրագրերի համար կիրառելի բառային հենք, որը ոչ միայն իր կազմության սկզբունքներով, այլև բառամիավորների ներկայացման եղանակներով տարբերվում է տպագիր բառարաններից: Ահա այս մոտեցմամբ այդ բառարանները, որոնցում տրվում է բառի կարճ և ամենաընդհանուր

⁵ Տե՛ս http://en.wikipedia.org/wiki/Svenska_Akademiens_ordbok Akademien:

իմաստի բացատրությունը, որևէ կերպ լիակատար էլեկտրոնային բառարան համարվել չեն կարող: Այնուամենայնիվ, փորձել ենք դրանք համեմատել օտարալեզու էլեկտրոնային բառարանների հետ. ի հայտ են գալիս նկատելի տարբերություններ: Այսպես՝ հայերեն էլեկտրոնային բառարաններն աչքի են ընկնում բառամիավորի խիստ միակողմանի ներկայացմամբ: Բնչպես հայտի է, օտարալեզու (հատկապես անգլերենի համար ստեղծված) էլեկտրոնային բառարաններում տրվում է գլխաբառի արտասանությունը, հայկականներում՝ ոչ, օտարալեզու բառարաններում արտացոլվում են գլխաբառի բոլոր իմաստները, հայկականներում տրվում է գլխաբառի մեկերկու ամենահիմնական իմաստների բացատրությունը, օտարալեզու բառարաններում յուրաքանչյուր իմաստի համար բերվում են օրինակ-նախադասություններ, հայկականներում՝ խիստ հազվադեպ, օտարալեզու բառարաններում ներկայացվում են բառի հոմանիշները, հակահիշները, այդ բառով կազմված դարձվածքները, տրվում է բառի ստուգաբանությունը, հայկականներում դարձվածքներն ու արտահայտությունները հիմնականում չեն ներկայացվում, հոմանիշներն ու հակահիշները՝ սակավ դեպքերում, իսկ բառի ստուգաբանությունն առհասարակ չի տրվում:

Համեմատաբար մեծ թիվ են կազմում թարգմանական բառարանները: Հայերեն թարգմանական բառարանները օտարալեզու բառարանների հետ համեմատելիս նկատում ենք մի շարք տարբերություններ: Լեզվական առումով շատ դժվար է սրանք լիակատար թարգմանական բառարաններ համարել, որովհետև թարգմանությունների ժամանակ հիմնականում տրվում է մուտքագրված բառի թարգմանական ընդամենը մեկ համարժեք՝ առանց օրինակ-նախադասությունների: Այս բառարանների թերություններից է նաև այն, որ չի տրվում թարգմանված բառի արտասանությունը: Սա հատկապես վերաբերում է անգլերեն-հայերեն և հայերեն-անգլերեն բառարաններին: Այն դեպքում, երբ նույնիսկ անգլերենի բառարաններում է տրվում այդ լեզվի բառերի արտասանությունը՝ հաշվի առնելով բառերի գրության ու արտասանության միջև եղած տարբերությունները, թարգմանական բառարաններում թարգմանված անգլերեն բառի արտասանությունը չներկայացնելը լրջագույն բացթողում է: Փաստորեն այս բառարանները կարելի է օգտագործել ընդամենը բառի ամենաառաջին և հիմնական իմաստն իմանալու համար:

Ընդհանրացնելով այսօր համցանցում հասանելի հայերեն էլեկտրոնային բառարանների ընդհանուր քննությունը՝ եզրակացրել ենք, որ դրանք լեզվաբանական տեսակետից խիստ ցածրորակ են: Իհարկե, նկատի ունենալով հայերենի լեզվական առնձնահատկությունները (բառապաշարի հարստություն, բառերի բազմիմաստություն, քերականական կարգերի առատություն ու զուգաձևություն) և համակարգչային բառարանագրության ոլորտում ունեցած ոչ մեծ փորձը՝ պետք է ասել, որ դժվար կլինի կազմել հայերենի էլեկտրոնային լիակատար բառարան, որում բառամիավորները ամբողջական կերպով կներկայացվեն, բայց դա ոչ միայն առավել քան անհրաժեշտ, այլև հնարավոր է: Կարևորել ենք էլեկտրոնային բառարաններ կազմելու աշխատանքները սիբիրոդական մակարդակից մասնագիտական ոլորտ տեղափոխելու անհրաժեշտությունը՝ ընդգծելով, որ դրանց կազմությամբ պետք է զբաղվեն համապատասխան տեխնիկա-

կան հմտություններ ունեցող լեզվաբան-բառարանագիրներ, որոնք ոչ միայն լեզվաբանական հարուստ գիտելիքներ կունենան, կտիրապետեն բառարաններ կազմելու դժվարին արվեստին, այլև քաջատեղյակ կլինեն ժամանակակից տեղեկատվական տեխնոլոգիաների նորագույն նվաճումներին և կկարողանան լեզվական տեղեկությունները տեխնիկական համապատասխան միջոցներով ներկայացնել:

Բառային ամբողջական տեղեկատվությունը համակարգչի միջոցով ուսումնասիրելու հեռանկարը համակարգչային բառարանագիրներին մղեց մշակելու դրանց ներկայացնելու բոլորովին նոր եղանակներ: Էլեկտրոնային բառարաններ կազմելու ընթացքում նրանք նկատեցին, որ էլեկտրոնային բառարաններում սրվում է բառամիավորների նկարագրությունը, մինչդեռ լեզվի բառապաշարի իմաստային-հասակացական կողմը որևէ կերպ չի արտացոլվում: Ուստի բառամիավորները նկարագրելուց բացի՝ փորձեր արվեցին ստեղծելու բառապաշարի այնպիսի հենք, որտեղ լեզվի բառապաշարը կբաժանվեր իմաստային-տրամաբանական խմբերի, հարկ եղած չափով կներկայացվեր բառապաշարի իմաստային կառուցվածքը. ծնվեց բառային հենքեր կազմելու գաղափարը: Այսօր գոյություն ունեցող բառային հենքերից իր կազմության սկզբունքներով և լայն տարածվածությամբ մյուսների համար օրինակ է ծառայում անգլերենի բառային հենքը՝ **WordNet**-ը, որը ստեղծվել է 1985թ. պրոֆեսոր Ջորջ Միլլերի կողմից Փրինսթոնի համալսարանում: 2012 թվականին ներկայացվեց WordNet-ի առցանց վերջին տարբերակը՝ **WordNet 3.1**-ը: Այն պարունակում է 155 287 բառ, որոնք ներկայացված են 117 659 հումանշային շարքերով և արտահայտում են 206 941 բառիմաստ:**WordNet**-ում մուտքագրված են միայն առարկա (գոյական), գործողություն (բայ) և հատկանիշ (ածական, մակբայ) ցույց տվող բառերը: Բաց են թողնված քերականական և վերաբերմունքային իմաստ արտահայտող բառերը: WordNet-ի մամուլությամբ ստեղծվել և շարունակում են ստեղծվել բառային հենքեր ինչպես որևէ առանձին լեզվի, այնպես էլ միանգամից մի քանի լեզուների համար (CWN (չինական Wordnet՝ 中文詞彙網路), WOLF (WordNet-ի ֆրանսերեն տարբերակը՝ WordNet Libre du Français), IndoWordNet, MultiWordNet և այլն)⁶: Աշխարհի մի քանի տասնյակ լեզուների համար արդեն ստեղծվել են բառային հենքեր, որոնք **The Global WordNet Associaton**-ի մաս են կազմում:

Առավել հաճախադեպ գործածվող բառային հենքերն իրենց հիմնական հատկանիշներով ներկայացնելուց հետո անդրադարձել ենք բառային հենքերի և տպագիր ու էլեկտրոնային բառարանների միջև եղած ընդհանրություններին ու տարբերություններին:

Ըստ էության՝ բառային հենքերի և էլեկտրոնային բառարանների միջ եղած մեծագույն տարբերությունը վերաբերում է դրանց կազմությանն ու դրանցում բառամիավորների նկարագրությանը: Այսպես՝ գիտենք, որ լեզվում նյութական իմաստ արտահայտող բառերը ցույց են տալիս հասկացություններ, որոնք էլ իրենց հերթին իմաստա-

⁶ Տե՛ս՝ www.globalwordnet.org:

յին զանազան հարաբերությունների մեջ են այլ հասկացությունների հետ: Բառային հենքեր կազմողները, հաշվի առնելով այս հանգամանքը, ներկայացրել են ոչ թե լեզվի բառերը, այլ հասկացությունները: Այստեղից էլ առաջ է գալիս էլեկտրոնային բառարանների և բառային հենքերի այն տարբերությունը, որ հենքերում կան իմաստային-հասկացական այնպիսի հարաբերություններ, որոնք էլեկտրոնային բառարաններում չկան: Մի բառը չի կարող մեկ այլ բառի տեսակ լինել, իսկ մի հասկացությունը կարող է մեկ այլ հասկացության տեսակ, մաս լինել: Այստեղից էլ տրամաբանորեն բխում է երկրորդ տարբերությունը. քանի որ հենքում ներկայացվում են հասկացությունները, ուստի ներկայացվում են նաև այդ հասկացությունների միջև եղած իմաստային հարաբերությունները, իսկ քանի որ էլեկտրոնային բառարաններում ներկայացվում են բառերը, հետևաբար տրվում են միայն բառիմաստային հարաբերությունները:

Կազմության եղանակի այս տարբերություններից էլ տրամաբանորեն առաջ են գալիս բառամիավորների նկարագրության սկզբունքների տարբերությունները: Ներկայացնելով բառամիավորների նկարագրության տարբերությունները սպագիր և էլեկտրոնային բառարանների միջև՝ ընդգծել ենք, որ ոչ էլեկտրոնային բառարաններում, ո՛չ բառային հենքերում բառի լիարժեք նկարագրությունը չի տրվում. բառային հենքերում կարևորվում են իմաստային-հասկացական հարաբերությունները, էլեկտրոնային բառարաններում՝ բառիմաստի նկարագրությունը: Մինչդեռ բառամիավորի ամբողջական նկարագրության համար անհրաժեշտ է տալ դրա հնչյունական, իմաստային և քերականական համակողմանի բնութագիրը, ուստի դրանցից որևէ մեկը չի կարող ուղենիշ ծառայել հայերենի համակարգչային շտեմարանի համար:

Ատենախոսության երկրորդ՝ **«Ժամանակակից հայոց լեզվի բառապաշարի համակարգչային շտեմարանի կազմության սկզբունքները»**, գլխում առաջարկվում են հայերենի բառապաշարի համակարգչային շտեմարանի կազմության սկզբունքները: Ամենից առաջ նշվում է, որ եղած էլեկտրոնային բառարաններից կամ բառային հենքերից որևէ մեկը որպես ուղենիշ չի կարող հանդես գալ հայերենի բառապաշարի շտեմարան կազմելու գործում, քանի որ դրանք նախ՝ կազմվել են այլ լեզուների համար՝ հատկապես անգլերենի, իսկ հայերենը, ընդհանրություններով՝ հանդերձ, նկատելի տարբերություններ ունի անգլերենի համեմատությամբ, հետո՝ ամենակարևորն այն է, որ հայերենի բառապաշարի համակարգչային շտեմարանը լիարժեքորեն համապատասխանի իր անվանը: Մեր շտեմարանում պետք է ոչ թե կարևորվի բառապաշարի իմաստային դասակարգումը և աչքաթող արվի բառիմաստների նկարագրությունը (ինչպես բառային մյուս հենքերում) և ոչ էլ պետք է առաջնությունը տրվի բառիմաստների նկարագրությանը, իսկ այդ բառիմաստների նշանակած հասկացությունների միջև եղած հարաբերությունները որևէ կերպ չներկայացվեն (ինչպես էլեկտրոնային բառարաններում): Հետևաբար, հայերենի շտեմարան կազմելիս պետք է մշակել բառի՝ իբրև լեզվի իմաստային հիմնական միավորի, համակողմանի ու սպառիչ նկարագրության նոր սկզբունքներ ու եղանակներ:

Շտեմարանի կազմության ուղղությամբ տարվող աշխատանքներն առաջարկվել է պայմանականորեն մի քանի փուլի բաժանել: Քանի որ շտեմարան կազմելը նման է իմաստային բարդ ցանց կազմելուն, ցանց, որտեղ բառերը միմյանց են կապված իմաստային մի շարք հարաբերություններով, հետևաբար խիստ կարևոր է կատարել շտեմարանում ներկայացվելիք բառացանկի ընտրություն: Թեպետ մինչև այժմ տարբեր լեզուների համար ստեղծված բառային հենքերում այդ լեզուների բառապաշարն ամբողջությամբ ներկայացված չէ, կարևորել ենք, որ մեր շտեմարանում ներկայացվեն և՛ առարկա (գոյական), և՛ գործողություն (բայ), և՛ առարկայի, գործողության հատկանիշ ցույց տվող բառերը (ածական, թվական, մակբայ), և՛ քերականական իմաստ արտահայտող բառերը (կապ, շղկապ), և՛ վերաբերմունքային իմաստ արտահայտող խոսքի մասերը (վերաբերական, ձայնարկություն):

Շտեմարանը պետք է ներկայացնի գրական արևելահայերենի բառապաշարի համաժամանակյա վիճակը: Սա ենթադրում է, որ շտեմարան չպետք է մուտքագրվեն հնացած բառերը, որոնք դուրս են եկել գործածությունից, քանի որ հասարակական կյանքում դրանց արտահայտած առարկաներն այլևս գործածական չեն, լեզվի հասարակական տարբերակներից՝ ժարգոնային բառերն ու արտահայտությունները, տարածական տարբերակները բարբառային բառերը: Փոխարենը շտեմարան պետք է մուտքագրվեն լեզվում ստեղծված ցանկացած նոր բառ և արտահայտություն կամ այն բառերը, որոնք վերաիմաստավորվել են:

Այնուհետև պետք է պարզել, թե շտեմարանում բառիմաստների բացատրությունը ըստ որ բացատրական բառարանի պետք է ներկայացվի: Առաջարկել ենք շտեմարանի բառերի բացատրությունը տալ ըստ 2014թ. Լ. Հովսեփյանի խմբագրությամբ լույս տեսած Էդուարդ Աղայանի «Արդի հայերենի առձեռն բացատրական բառարանի», որը հեղինակի «Արդի հայերենի բացատրական բառարանի» վերամշակված տարբերակն է և կազմության սկզբունքներով համապատասխանում է բառապաշարի ընտրության վերը նշված սկզբունքներին:

Հաջորդ քայլը հայերենի բառապաշարի իմաստային դասակարգում կատարելն է: Սա, թերևս, ամենալուրջ և ամենադժվար գործն է: Դա անելու երկու եղանակ կա: Առաջին եղանակով կարող ենք մշակել համապատասխան սկզբունքներ, որոնց միջոցով կկատարվի հայերենի բառապաշարի իմաստային դասակարգում (բառապաշարը կբաժանվի իմաստային-տրամաբանական խմբերի՝ ըստ հասկացական տարբեր դաշտերի): Միանգամից նկատենք, որ այդպիսի դասակարգում կատարելը շատ բարդ խնդիր է. դա անելու համար անհրաժեշտ է ունենալ վերլուծող, դասակարգող ու ընդհանրացնող միտք, բացի դրանից՝ հրաշալի պատկերացնել հայերենի բառապաշարի իմաստային կառուցվածքը: Այնուհետև, առկա բառային հենքերում դեռևս ոչ մի լեզվի համար բառապաշարի օրինակելի դասակարգում չի կատարվել, հետևաբար որևէ երաշխիք չկա, որ մենք հաջողության կհասնենք: Բառապաշարի իմաստային դասակարգման երկրորդ եղանակը նկատելիորեն կտարբերվի առաջինից: Ինչպես նշեցինք, WordNet-ում կատարվել է անգլերենի բառապաշարի իմաստային դասակարգում, բառապաշարի առանձնացումն ըստ հասկացական դաշտերի կատարվել է նաև

համընդհանուր ցանցային լեզվում (Universal Networking Language)⁷: Առաջարկել ենք հայերենի բառապաշարը նախապես բաժանել իմաստային խմբերի՝ ըստ WordNet-ի հասկացական դաշտերի, հետո, երբ ավելի քիչ թվով բառերի իմաստային քննությունից հայերենի բառապաշարի իմաստային կառուցվածքն ավելի տեսանելի կլինի, իմաստային առանձնահատկությունները նկատելի, հնարավոր կլինի մշակել հայերենի բառապաշարի իմաստային դասակարգման ընդհանրական սկզբունքներ:

Այնուհետև ներկայացվել է WordNet բառային հենքն իր բնութագրական ամենահիմնական հատկանիշներով՝ կազմության սկզբունքներ, կառուցվածքային առանձնահատկություններ, առանձնացվել են իմաստային-քերականական տարբեր խմբերի միջև գործող հասկացական տարբեր հարաբերություններն ու դրանց ներկայացման եղանակները: Քննելով մեր շտեմարանի կազմության համար բառային այս հենքի օրինակելի լինելու հարցը՝ ընդգծել ենք, որ առանց բովանդակային լուրջ փոփոխությունների՝ այն չի կարող դառնալ հայերենի համակարգչային շտեմարանի հիմք: Օրինակ խիստ թերի է, ասենք, գոյականների բաժանումը հասկացական 25 դաշտերի: Դժվար է ասել՝ այդ հասկացական դաշտերում գոյականների միջև գործող իմաստային հարաբերությունները արդյո՞ք միայն հիպերնիմություն-հիպոնիմություն և մերոնիմություն-հոլոնիմություն զույգերն են, և էթե անգամ այո՞, ո՞րչափով դրանք կարտահայտեն հայերենի գոյականների անվանական-հասկացական հարաբերությունները: Նույն կերպ էլ, նկատի ունենալով բայ խոսքի մասի բազմիմաստությունն ու ձևառատությունը, խիստ կասկածելի է բայիմաստների միջև միայն հիպերնիմային, տրոպոնիմային և իմաստային հանգեցման հարաբերություններ առանձնացնելը: Հետևաբար, օգտվելով այս բառային հենքում եղած լեզվական նյութից՝ հայերենի բառապաշարի իմաստային դասակարգումը պետք է կատարել բոլորովին նոր սկզբունքներով ու եղանակներով: Եվ էթե նույնիսկ հայերենի բառապաշարի իմաստային դասակարգումը նկատելի տարբերություններ հանդես բերի WordNet-ից և UNL-ից, մինևույն է, կկարողանանք մեր շտեմարանը կիրառելի դարձնել UNL համակարգում և փոխգործակցելի WordNet-ի հետ: Բարեբախտաբար, դրանցում կարևորվում է հասկացությունների ներկայացումը, իսկ հասկացությունները բոլոր լեզուներում էլ նույնն են: Տարբեր լեզուներում տարբեր են միայն այդ հասկացություններն անվանող բառերը: Ուստի մեր շտեմարանը գուցե կտարբերվի բառապաշարի՝ ըստ հասկացական տարբեր դաշտերի առավել լիակատար ներկայացմամբ ու նկարագրությամբ, բայց ոչ տրամաբանական անհամատեղելիությամբ:

Ատենախոսության վերջին՝ **«Ժամանակակից հայոց լեզվի բառապաշարի հասկարգչային շտեմարանի բառահոդվածների կառուցվածքը»**, գլխում ներկայացվում են նոր կազմվելիք շտեմարանում բառամիավորների նկարագրության սկզբունքները, որոնցով փորձ է արվում տալ լեզվի իմաստային հիմնական միավորի՝ բառի հնարա-

⁷ **Ուչիդա Հ., Զու Մ., Դելլա Սենտա Ս.**, Համընդհանուր ցանցային լեզու, Եր., «Գիտություն» հրատ., 2007, էջ 5:

վորինս ամբողջական ձևախմբաբանական բնութագիրը: Առաջարկվել է բառամիավորները նկարագրել հետևյալ դաշտերով:

1. Հնչյունական

Այս դաշտում տրվում է բառի արտաքին հնչյունական կողմի նկարագրությունը, արտասանությունը, գիտական տառադարձումը, հնչյունափոխությունը, վանկային կառուցվածքը:

Բառային հենքերում, ճիշտ է, բառի արտասանությունը չի ներկայացվում: Հայերեն սպագիր և էլեկտրոնային բառարաններում էլ, հաշվի առնելով գրության և արտասանության միջև եղած ոչ մեծ տարբերությունները, բառերի արտասանությունը չի տրվում: Մակայն կարծում ենք, որ մեր շտեմարանում բառի ուղղագրությանը զուգահեռ ճիշտ կլինի ներկայացնել նաև դրա արտասանությունը: Ինչ խոսք, մեզ արևելահայերիս համար, հայերեն բառերն արտասանելն առանձնակի դժվարություն չի ներկայացնում, բայց էթե հաշվի առնենք այն հանգամանքը, որ այդ շտեմարանից օգտվողների մի մասը պատմական դեպքերի բերումով ապրում է Հայաստանից դուրս, հայերեն քիչ գիտի կամ ընդհանրապես չգիտի, և նրանց համար հայերեն սովորելը նման է օտար լեզու սովորելուն, իսկ հայերեն բառեր սովորելը՝ օտար լեզվի բառեր սովորելուն (չխոսելով հայերենով հետաքրքրվող օտարազգիների մասին), շտեմարանը նրանց համար ոչ միայն բառի մասին ամբողջական, լեզվական ու հանրագիտարանային բնույթի տեղեկատվություն ստանալու, այլև այդ բառի ճիշտ գրական արևելահայերեն արտասանությունը իմանալու բացառիկ հնարավորություն է: Բացի դրանից՝ սա կարևոր է նաև բառարանի լիակատարության տեսակետից, քանի որ տեղեկություններ կտրվեն ոչ միայն ուղղագրական, այլև ուղղախոսական նորմերի մասին:

2. Բառագիտական

ա. Իմաստաբանական. ամենից առաջ կներկայացվեն բառի իմաստային բնութագրումներն ու կապը այլ բառերի հետ (բառիմաստի բացատրություն, ուղղակի և փոխաբերական իմաստներ, հոմանիշներ, հակահիշներ, համանուններ), ապա՝ հասկացական հարաբերությունները՝ հիպերնիմություն, հիպոնիմություն, հոլոնիմություն, տրոպոնիմություն են:

բ. Բառակազմական. կներկայացվի բառի ձևային կազմը, կառանձնացվեն բառի արմատներն ու ածանցները, բառակազմական և հիմքակազմական այլ տարբերը: Նկատենք, որ ոչ սպագիր ու էլեկտրոնային բառարաններում և ոչ ուրիշ լեզուների համար մինչև այժմ կազմված բառային հենքերում բառի բառակազմական բնութագիրը չի ներկայացվում: Մեր շտեմարանում, սակայն, բառի բառագիտական լիակատար բնութագիրը ներկայացնելու համար ճիշտ կլինի տալ նաև դրա կազմությունը:

գ. Ստուգաբանական. կտրվի բառի ստուգաբանությունը: Տպագիր բառարաններում, ինչպես հայտնի է, բառի ստուգաբանությունը բառիմաստի բացատրությանը զուգահեռ սովորաբար չի ներկայացվում: Հայերեն էլեկտրոնային բառարաններում ևս որոնված բառի ստուգաբանությունը չի տրվում, ուստի հայերեն որևէ բառի ստուգաբանությունն իմանալու միակ եղանակն այն Հրաչյա Աճառյանի «Հայերեն արմա-

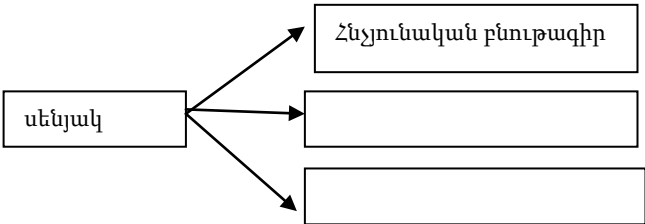
տական բառարանում» փնտրելն է: Օտարալեզու էլեկտրոնային բառարանների հիմնական մասում, մինչդեռ, բառիմաստի բացատրությունից, հոմանիշների, հականիշների ներկայացումից բացի՝ տրվում է նաև այդ բառի ստուգաբանությունը: Որքան էլ զարմանալի է, բառային հենքերում ևս բառի ստուգաբանությունը չի ներկայացվում: Սակայն բառը լիակատար ձևով ներկայացնելու համար անհրաժեշտ է տալ նաև դրա ստուգաբանությունը, որը բառի լիակատար նկարագրության անհրաժեշտ տարրերից մեկն է:

3. Քերականական

Բառի քերականական հատկանիշների ներկայացման ժամանակ հատկապես կկարևորվի ձևաբանական բնութագիրը (կտրվի բառաձևի խոսքիմասային պատկանելությունը, այդ խոսքի մասին հատուկ քերականական կարգերը): Ներկայիս բառային հենքերում խիստ թերի է կատարվում բառաձևի նկարագրությունը: Օրինակ՝ որևէ կերպ չեն ներկայացվում բայաձևերը, նույնիսկ էթե որևէ ժամանակաձևով բայ մուտքագրենք, որոնման արդյունքում ոչինչ չի հայտնվի: Այս առումով նույնիսկ տպագիր բառարաններն ավելի նախընտրելի են, քան բառային հենքերը: Օրինակ էթե անգլերենի WordNet մուտքագրենք to eat «ուտել» բայի անցյալ անորոշի ձևը՝ ate-ը, որոնման արդյունքում որևէ բան չենք գտնի, իսկ այդ բառը անգլերենի տպագիր ցանկացած բառարանում փնտրելիս ոչ միայն կիմանանք, որ դա to eat «ուտել»-ի անցյալի ձևն է, այլև հղում կտրվի to eat բային: Այս հանգամանքը հաշվի առնելով էլ՝ կարծում ենք, որ հայերենի համակարգչային շտեմարանը լիակատար կլինի, էթե դրանում բառաձևերն էլ բառամիավորների պես ամբողջական ու համակողմանիորեն ներկայացվեն:

Այս ամենից բացի՝ առաջարկվել է շտեմարանը վերածել նաև թարգմանական բառարանի: Կարելի է կամ որոնված բառի թարգմանության համար հղում տալ համապատասխան լեզվի բառային հենքին, ընդ որում՝ ոչ միայն բառի, այլև մասնավոր բառիմաստի դեպքում, կամ հղում տալ այդ լեզվի թարգմանական առցանց լավագույն բառարաններին:

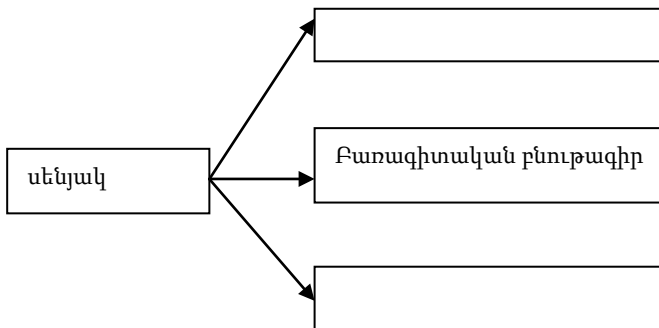
Բառամիավորների ներկայացման եղանակները ներկայացնելուց հետո բերվում է շտեմարանում ներկայացվելիք բառահոդվածի օրինակելի նմուշ: Այսպես՝ էթե համակարգչային շտեմարան մուտքագրենք **սենյակ** բառը՝ և ցանկանանք իմանալ դրա հնչյունական բնութագիրը, պատկերն այսպիսին կլինի՝



Մենյակ

- **Հ:** (գոյ) սենյակ
- Արտասանություն
- **Հնչյունական կազմ**
- **Ա**
- Չայնավոր
- **Ըստ բերանում լեզվի գրաված դիրքի**
- Հետնալեզվային
- **Ըստ բերանում լեզվի բարձրացման աստիճանի**
- Ստորին բարձրացում
- **Ըստ շրթայնացած կամ ոչ շրթայնացած լինելու հատկանիշի**
- Ոչ շրթային
- **Հնչյունափոխություն**
-
- **Վանկային կառուցվածք**
- սեն (փակ) + յակ (փակ)
- **Տառադարձում**
- Senyak

Իմաստային բնութագիրը ներկայացնելիս կտրվեն բառիմաստի բացատրությունը, իմաստային-հակացական հարաբերությունները՝



Մենյակ

- **Հ: (գոյ) սենյակ**, սրահ, դահլիճ (բնակարանի կամ շինության յուրաքանչյուր բաժանմունք, որ մյուսներից բաժանվում է պատերով) {Նա դուրս եկավ սենյակից}
- **Հականիշ**
- -----
- **Արտահայտություն**
- **Հ: (գոյ)** սենյակի երկարությունը չափել (սենյակում քայլել)
- **Ուղիղ հիպոնիմ / հիպոնիմային շարք**
- **Հ: (գոյ)** պալատ (թագավորի բնակության շքեղ շենք)
- **Հ: (գոյ)** նախասենյակ (շենքի մուտքի առջև կառուցված փոքրիկ շինվածք)
- **Հ: (գոյ)** գարդասենյակ (դերասանների հարդարվելու և զգեստավորվելու սենյակը)
- **Հ: (գոյ)** խոհանոց (վառարանով, օջախով, սալիկով ևն հարմարություններով շենք կամ սենյակ)
- **Հ: (գոյ)** ընդունարան (շենքի, բնակարանի մեջ սենյակ կամ սրահ որտեղ ընդունում են այցելուներին)
- **Հ: (գոյ)** հյուրասենյակ (հյուր ընդունելու համար նախատեսված սենյակ)
- **Հ: (գոյ)** աշխատասենյակ (աշխատանք կատարելու համար հատկացված սենյակ)
- **Հ: (գոյ)** ներքնագավիթ (ներքին գավիթ)
- **Հ: (գոյ)** ճաշասենյակ (ճաշելու սենյակ)
- **Հ: (գոյ)** ննջասենյակ (քնելու սենյակ, տան այն սենյակը, որտեղ քնում են)
- **Հ: (գոյ)** լոգասենյակ (լոգանք ընդունելու սենյակ)
-
- **Ուղիղ հիպերնիմ / հիպերնիմային շարք**
- **Հ: (գոյ)** տարածք (տարածություն, մակերես)
- **Ուղիղ հիպերնիմ / հիպերնիմային շարք**
- **Հ: (գոյ)** տարածք (տարածություն, մակերես)
 - **Հ: (գոյ)** կառույց, շինություն, շենք
 - **Հ: (գոյ)** արհեստածին իրեր
 - **Հ: (գոյ)** հասկացություն

Մերոնիմ

- **Հ: (գոյ)** առաստաղ (շինության ծածկի ներսի մասը)
- **Հ: (գոյ)** դուռ (ելումուտ անելու բացվածք պատի, պարսպի և այլնի մեջ)
- **Հ: (գոյ)** պատ (շինության կառույցի ուղղահայաց մասը, որ ծառայում է որպես ծածկի հենարան և շենքը մասերի միջնորմ)
- **Հ: (գոյ)** պատուհան (օդ և լույս ստանալու համար պատի մեջ բացված անցք)

Հոլոնիմ

- **Հ: (գոյ)** շենք, շինություն
Բառակազմությունը`
- **Հ: (գոյ) սենյակ**
- **Կազմություն**
- պարզ (բառեր, որոնք կազմված են բառակազմական մեկ հիմնական ձևությից)

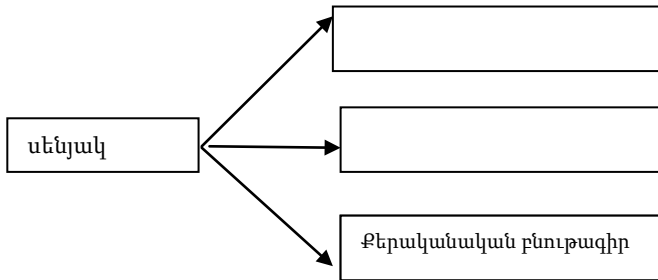
Ստուգաբանությունը`

- **Հ: (գոյ) սենյակ**
- **Ստուգաբանություն**
- Անհայտ

Քերականական բնութագրի ժամանակ կներկայացվեն **սենյակ** բառաձևի քերականական կարգերը`

Հ: (գոյ) սենյակ

- **Քերականական կարգ**
- **Հոլով**
- Ուղղական
- **Թիվ**
- Եզակի
 - **Հ: (գոյ)** գոյական (առարկա ցույց տվող բառեր)
 - **Հ: (գոյ)** լիմաստ բառ (նյութական աշխարհի առարկաներ անվանող բառեր)
 - Հ: (գոյ)** բառ (իմաստ և ինքնուրույն գործածություն ունեցող հնչյունախումբ, որն ունի բառային շեշտ)



Աշխատանքն ավարտվում է հետազոտության արդյունքներն ամփոփող հետևյալ եզրակացություններով:

1. Նախորդ դարի 50-ական թվականներից սկսած, լեզվաբանության մեջ համակարգչային տեխնիկայի տարաբնույթ կիրառությամբ պայմանավորված՝ ձևավորվեց նոր գիտակարգ՝ *համակարգչային լեզվաբանությունը*:

2. Համակարգիչների կիրառությունը լեզվաբանության մեջ ունի երկակի բնույթ. մի կողմից՝ լեզվի միջոցով է հասանելի դառնում եղած հսկայական տեղեկատվությունը, մյուս կողմից՝ համակարգիչների անմիջական մասնակցությամբ է ուսումնասիրվում, համակարգվում և ներկայացվում լեզվական նյութը:

3. Լեզվաբանության մեջ համակարգիչների կիրառությունը հնարավոր դարձրեց հաջող կերպով (թեև ոչ լիակատար ձևով) հաղթահարել լեզվական պատմեշները մեքենական թարգմանության ոլորտում, ստեղծել արհեստական բանականության համակարգեր: Համակարգչային լեզվաբանության հետագա զարգացումը, լեզվի (այդ թվում իմաստի) ձևայնացման ու համակարգչի միջոցով ներկայացնելու ուղիներ փնտրելու ընթացքում ձևավորվեց *համակարգչային բառարանագրությունը*, որն այսօր, թերևս, համակարգչային լեզվաբանության ամենաարագ զարգացող բնագավառներից մեկն է:

4. Համակարգչային բառարանագրության նպատակը բառի թե՛ հասկացական, թե՛ լեզվական իմաստների ներկայացումն է համակարգչի միջոցով: Ներկայումս բառի հասկացական և լեզվական իմաստների ներկայացման երկու եղանակ կա՝ էլեկտրոնային բառարանների և բառային հենքերի միջոցով:

5. Տեղեկատվական տեխնոլոգիաների կիրառությունը լեզվաբանության մեջ լեզվական տեսակետից զարգացման նոր աստիճանի հասցրեց համակարգչային բառարանագրության շրջանակներում տարվող աշխատանքները: Աշխարհի բազմաթիվ լեզուների համար ստեղծվել են միալեզվային և բազմալեզվյան (թարգմանական) բառարաններ, որոնք ունեն կազմության սկզբունքների և բառամիավորների ներկայացման եղանակների ընդհանրություններ: Էլեկտրոնային բառարանները լեզվական տեսակետից ունեցան աստիճանական զարգացում. սկզբնական շրջանի էլեկտրոնային բառարանները տպագիր բառարանների թվային տարբերակներն էին, այնուհետև

ձևով ու բովանդակությամբ նոր բնույթի էլեկտրոնային բառարաններ ստեղծելու փորձեր արվեցին: Այս նոր տեսակի էլեկտրոնային բառարաններն աչքի են ընկնում լեզվի բառապաշարի լայն ընդգրկմամբ և բառամիավորների առավել համակողմանի նկարագրությամբ:

6. Հարուստ պատմություն ունեցող հայ բառարանագրությունն էլ փորձեց անմասն չմնալ և առաջին քայլերը կատարեց էլեկտրոնային բառարաններ կազմելու ուղղությամբ, սակայն այս բնագավառում մեր համակարգչային բառարանագիրների կողմից տարվող աշխատանքները չի կարելի գոհացուցիչ համարել:

7. Քանի որ էլեկտրոնային բառարաններում տրվում է միայն բառիմաստը, սովյալ բառի նշանակած հասկացությունը ոչ լիարժեք կերպով է ներկայացվում, իսկ այդ և մյուս հասկացությունների միջև եղած հարաբերությունները գրեթե ցույց չեն տրվում, համակարգչային բառարանագիրները ձեռնամուխ եղան *բառային հենքերի* կազմությանը:

8. Բառային հենքերում, ի տարբերություն էլեկտրոնային բառարանների, ավելի շատ կարևորվում է բառի հասկացական կողմի ներկայացումը: Ըստ այդմ՝ բառապաշարը նախ՝ բաժանվում է իմաստային-քերականական խմբերի (խոսքի մասեր), ապա՝ դրանց ներսում՝ հասկացական-տրամաբանական խմբերի, այնուհետև առանձնացվում ու ներկայացվում են հասկացական և իմաստային հարաբերությունները: Բառային հենքեր ստեղծվել են աշխարհի տարբեր լեզուների համար. դրանց հիմքում ընկած է անգլերենի բառային հենքը՝ WordNet-ը: WordNet-ում ներկայացված են միայն առարկա (գոյական), գործողություն (բայ), հատկանիշ (ածական, թվական, մակբայ) ցույց տվող բառերը:

9. Հայերենի համակարգչային շտեմարանում հայերենի բառապաշարի ներկայացման համար կարևորվեց այն հանագամանքը, որ դրանում պետք է արտացոլվի արդի արևելահայերենի համաժամանակյա վիճակը, այսինքն՝ ճիշտ կլինի համակարգչային շտեմարան ներմուծել միայն գրական արևելահայերենի գործուն բառապաշարին պատկանող բառերը:

10. Համակարգչային շտեմարանում հայերենի բառապաշարն ամբողջական դարձնելու համար առաջարկվել է ընդգրկել ոչ միայն նյութական իմաստ արտահայտող խոսքի մասերը, այլև սպասարկու բառերն ու քերականական իմաստ արտահայտող խոսքի մասերը:

11. Քանի որ շտեմարանում ներկայացվելիք բառապաշարի ընտրությունից բացի՝ պետք է կատարել նաև իմաստային դասկարգում, առաջարկվել է այդ դասակարգումը կատարելիս նախապես օգտվել անգլերենի բառային հենքում՝ WordNet-ում, եղած դասակարգումից, այնուհետև, երբ հայերենի բառապաշարի իմաստային կառուցվածքն ավելի ակնհայտ կդառնա, կմշակվեն ընդհանրական այնպիսի սկզբունքներ, որոնցով կատարված դասակարգումը ոչ միայն լիարժեքորեն կներկայացնի հայերենի բառապաշարը, այլև ճշտորեն կարտացոլի իմաստային-քերականական տարբեր խմբերի միջև եղած հարաբերությունները:

12. Բառացանկի ընտրությունից և իմաստային դասակարգումից հետո պետք է տրվի համակարգչային շտեմարան մուտքագրված բառի լիակատար հնչյունական, բառագիտական և քերականական նկարագրությունը:

13. Հնչյունական բնութագրի համար առաջարկվում է տալ բառի արտասանությունը, բառի արտաքին հնչյունական կողմի նկարագրությունը, արտասանությունը, գիտական տառադարձումը, հնչյունափոխությունը, վանկային կառուցվածքը:

14. Իմաստային բնութագրի համար առաջարկվել է ներկայացնել բառն այդ իմաստով իր հոմանշային շարքով (բազմիմաստ բառերի դեպքում՝ ըստ տարբեր իմաստների), տալ փնտրված բառի իմաստի բացատրությունը, բերել օրինակ-նախադասություններ: Բառիմաստի բացատրությունից հետո ներկայացնել բառի ուղղակի և փոխաբերական իմաստները, համանունները, հականիշները, այդ բառով կազմված արտահայտություններն ու դարձվածքները: Իմաստային հարաբերությունները ներկայացնելուց հետո տալ հասկացական հարաբերությունները (հիպոնիմություն, հիպերնիմություն, հոլոնիմություն, տրոպոնիմություն են): Բառակազմական բնութագրի ժամանակ ներկայացնել բառի ձևային կազմը (առանձնացնել բառի արմատներն ու ածանցները): Ստուգաբանական բնութագրի ժամանակ ներկայացնել բառի ստուգաբանությունը:

15. Քերականական բնութագրի ժամանակ առաջարկեցինք ներկայացնել քերականական հատկանիշները՝ հատկապես կարևորելով ձևաբանական բնութագիրը (տալ բառաձևի խոսքիմասային պատկանելությունը, այդ խոսքի մասին հատուկ քերականական կարգերը):

Ատենախոսության հիմնադրույթներն արտացոլված են հետևյալ հրապարակումներում:

1. **Հակոբյան Ֆ.** Հայերենի բառապաշարի համակարգչային շտեմարանի կառուցման ուրվագիծ // Բանբեր Երևանի համալսարանի. Երևան, 2014, 144.2. (№ 3), էջ 60-70:
2. **Հակոբյան Ֆ.** Բառապաշարի համակարգչային շտեմարաններ // Երևանի Վ. Բրյուսովի անվան պետական լեզվահասարակագիտական համալսարանի Բանբեր. Երևան, 2014, 1 (30), էջ 131-143:
3. **Հակոբյան Ֆ.** Բառապաշարի ներկայացումը էլեկտրոնային բառարաններում // Խ. Արովյանի անվան հայկական պետական մանկավարժական համալսարան. Հայագիտական հանդես, Երևան, 2014, № 4 (28), էջ 107-116:
4. **Հակոբյան Ֆ.** Բառապաշարի ներկայացումը հայերեն էլեկտրոնային բառարաններում // «Կանթեղ» գիտական պարբերական, 2014, 2-3 (59-60), էջ 119-127:

АКОПЯН ФРИДА АРТАШЕСОВНА

**ПРИНЦИПЫ ОПИСАНИЯ СЛОВАРНЫХ ЕДИНИЦ В ЭЛЕКТРОННОЙ
ЛЕКСИЧЕСКОЙ БАЗЕ СОВРЕМЕННОГО АРМЯНСКОГО ЯЗЫКА**

Диссертация на соискание ученой степени кандидата филологических наук по специальности 10.02.02 – “Общее и сравнительное языкознание”.
Защита состоится 25 января 2016г. в 15⁰⁰ ч. на заседании специализированного совета ВАК по лингвистике 019 при Институте языка им. Р.Ачаряна НАН РА (адрес: 0015, Ереван, ул. Григора Лусаворича 15)

РЕЗЮМЕ

Целью диссертационной работы является разработка принципов построения нового типа электронного лексического банка данных современного армянского языка.

Научная новизна и практическое значение работы заключаются в том, что как процесс создания, так и результат построения принципиально нового банка данных будет способствовать дальнейшему расширению использования информационных технологий в области гуманитарных наук, изучению некоторых теоретических проблем компьютерной лингвистики, разработке методов формального описания языка, а также решению различных прикладных задач.

Диссертация состоит из введения, трех глав, заключения и списка использованной литературы.

Во введении рассматриваются вопросы, связанные с основными особенностями компьютерной лингвистики, стоящими перед ней теоретическими и прикладными задачами, в частности, в области применения методов формального описания языков в целях изучения и представления языковой структуры с помощью компьютерной техники, способами и приемами описания лексики, а также представляются цели и задачи исследования.

В первой главе- «Предпосылки создания электронного лексического банка данных современного армянского языка» -рассматриваются имеющиеся в интернете источники по предоставлению лексической информации, с целью выявления, могут ли они считаться электронной базой данных и послужить примером для построения электронного лексического банка для современного армянского языка. Описывается структура иноязычных электронных словарей и лексических баз, говорится об их общих чертах и различиях в представлении словарных единиц. Отмечается, что в них ни в коей мере не дается исчерпывающая формально-семантическая характеристика слова,

и посему они не только не могут служить прямой основой для построения нашего банка, но и сами по себе не могут считаться лексической базой данных. Обращается внимание на работы, ведущиеся в сфере армянской электронной лексикографии, рассматриваются вопросы о действительном существовании электронных словарей армянского языка и о том, в какой мере имеющиеся источники могут быть полезны для создания армянского банка данных. Подчеркивается важность представления в будущем банке данных армянской лексики в ее современном состоянии, с наиболее точной семантической классификацией представленной лексики - наряду с полным описанием фонетических, семантических и грамматических особенностей лексических единиц.

Во второй главе - «Принципы составления электронного лексического банка данных современного армянского языка»- предлагаются основополагающие принципы построения. Рассматривается вопрос о выборе лексических единиц, подчеркивается важность представления не только знаменательных, но также и служебных частей речи и модальных слов. Понятийно-логическую классификацию выбранной лексики по семантическим полям предполагается осуществить согласно с лексической базой данных WordNet для английского языка. При этом представлены основные характеристики WordNet-a.

В третьей главе - «Структура словарной статьи в лексическом банке данных современного армянского языка»- представляются принципы описания лексических единиц предполагаемого банка данных, согласно с которыми делается попытка представить по возможности полную формально-семантическую характеристику слова как основной смысловой единицы языка. Предложено представить фонетические, лексикологические и грамматические характеристики по соответствующим полям. Фонетическое поле будет отражать произношение, описание внешней фонетической стороны слова, его слоговую структуру, позиционные звукоизменения, транскрипцию и транслитерацию. Лексикологическое поле содержит семантические и понятийные взаимоотношения между словами, а грамматическое поле будет содержать информацию о грамматических категориях и частеречной принадлежности слов.

В заключении даются основные выводы исследования. Отмечается, что при наличии ограниченного опыта по составлению электронных словарей армянского языка создание полноценного лексического банка данных представляется сложной задачей, однако это не только крайне необходимое, но и осуществимое начинание. Полагаясь на основе предложенных в диссертационной работе принципов можно достичь наиболее полного формально-семантического описания лексических единиц современного армянского языка, опираясь на многовековые традиции рукописной и печатной армянской лексикографии, а также используя опыт составления современных электронных баз данных для других языков.

HAKOBYAN FRIDA ARTASHES

**(THE PRINCIPLES OF WORD DESCRIPTIONS IN THE LEXICAL
DATABASE OF MODERN ARMENIAN LANGUAGE)**

Thesis for the degree of candidate of philological sciences, specialty 10.02.02-
“General and Comparative Linguistics”.

The defense of the thesis will be held on 25 January, 2016, at 15⁰⁰, at the session of the Specialized Council on Linguistics 019 HAC (Higher Attestation Commission) in the Institute of Language after H. Acharyan of National Academy of Sciences, RA (address: 0015, st. G. Lusavorich, 15).

RESUME

The aim of the dissertation is on one hand by using current advances of modern computational lexicography, on the other hand, taking consideration linguistic features of the Armenian language to develop principles of creating a lexical database in which the lexicon will be presented in its current state, the semantic classification of the lexicon will be done, lexical and semantic relations of words will be shown and lexical-semantic characteristics of words will be given.

The novelty and practical significance of the thesis lies in developing principles of creating a kind of lexical database which will have an enormous impact on the way of using modern information technologies in the field of human science, investigating linguistic phenomena as well as will have many different applications.

The dissertation consists of an introduction, three chapters, conclusions and bibliography.

The introduction presents some issues connected to computational linguistics, formal description of language, types of meaning representations on the Internet. The choice of the theme under the study, scientific novelty, the modernity of the dissertation, the aims and the methods are discussed as well.

In the **first chapter** entitled “The background of creating a lexical database for the Armenian language” some of modern age electronic dictionaries and lexical databases compiled for different foreign languages are shown as the main source of presenting lexical information. Major characteristics and differences between electronic dictionaries and lexical databases are analyzed. It is noted that neither in electronic dictionaries nor lexical databases is given lexical-semantic characteristics of the word as the principal semantic unit of language; for us none of them can serve as an example for creating a lexical database. Moreover, they cannot be considered lexical databases by themselves in any way because neither of them gives complete lexical-semantic characteristics of words. Some issues about Armenian dictionaries are discussed trying to find ways how the experience of Armenian lexicography can be used in order to create a lexical database. It is noted that in our newly compiled database must be given phonetic, lexical-semantic and morphological characteristics of words.

The **second chapter** entitled “The suggested principles of building a lexical database for the Armenian language” begins with a brief overview of WordNet – a large lexical database of English by analyzing its structure, separating the most frequently encoded conceptual-semantic relations and then moves on its main focus-defining the guidelines and principles of the database compilation.

In the **third chapter** entitled “The structure of headwords in the lexical database for the Armenian language” are presented the principles according to which it is aimed to describe headwords of modern Armenian language giving formal as well as semantic characteristic of the word as the principal semantic unit of language by separating different fields.

The phonetic field will reflect pronunciation, external phonetic part of words as well as syllabic structure, transcription, transliteration. The semantic field consists of lexical and semantic descriptions of words, the morphological field will contain information about grammatical categories and morphological structure of words.

In the conclusions, the main findings and some inferences drawn from the research are summarized.

It is noted that creating a lexical database for the Armenian language is a difficult task because of lack enough experience in the area of modern electronic lexicography, but it is not impossible because Armenian lexicography has a long history and rich traditions. By using current advances of modern computational linguistics with the suggested principles we can create a lexical database in which the word as the principal semantic unit of language will be presented and its complete lexical-semantic characteristics will be given.