

ՄԵՔԵՆԱԿԱՆ ԹԱՐԳՄԱՆՈՒԹՅԱՆ ԽՆԴԻՐՆԵՐԸ

Մեքենական թարգմանության ծննդյան թիվը համարվում է 1947թ. մարտը, իսկ հեղինակները՝ Ուորեն Ուիվերը և Նորբերտ Վիները:

Էլեկտրոնային հաշվիչ մեքենաների հայտնվելուց հետո բնական լեզուների թարգմանության համար սկսեցին կիրառել համակարգիչներ:

ՄԹ սկզբնական համակարգերը ունեին միայն ընդարձակ երկլեզվյան բառարաններ, որտեղ տրված էին մուտքային լեզվի գլխաբառերը, իսկ դրանց դիմաց ելքային լեզվի համապատասխան բառերը, տրված էին նաև քերականական որոշ կանոններ, հիմնականում շարադասության վերաբերյալ: Շատ շուտով պարզ դարձավ, որ շարադասության մասին տրված կանոնները խիստ անձակ են, և կարիք կա շարահյուսական վերլուծության համակարգված մեթոդի: Ստեղծվեցին բազմաթիվ ծրագրեր՝ ժամանակակից լեզվաբանության հնարավորություններով, հատկապես ձևային քերականության օրինակներով կատարյալ թարգմանության համար: Ուսումնասիրության առաջին տասնամյակում մասնագետները լավատեսությամբ էին լցված, մինչ կհանդիպեին անլուծելի թվացող իմաստաբանական խնդիրների:

Նշված ժամանակահատվածում կային երկու համակարգեր՝ Մարկ II համակարգը, որը ստեղծվել էր IBM –ի և Վաշինգտոնի համալսարանի կողմից, և Ջորջթաունի համալսարանի ստեղծած համակարգը, սակայն այդ երկուսի արդյունքներն էլ հիասթափեցնող էին: Դրոնք հիմնականում վերաբերում էին քիմիային: 1964թ. ԱՄՆ ծրագրի հովանավորները սկսեցին մտահոգվել համակարգերի անհաջողությամբ: Նրանք հիմնեցին ALPAC-ը (Automatic Language Processing Advisory Committee), որը 1966 թ-ի գեկույցով եկավ այն եզրակացության, որ ՄԹ ավելի դանդաղ է, պակաս ճշգրիտ և կրկնակի թանկ, քան մարդու կատարած թարգմանությունը, և որ ՄԹ հեռանակար չունի, և հարկ չկա ներդրում անել մեքենական թարգմանության ոլորտում, և դրա փոխարեն նրանք առաջարկեցին ներդրումներ անել մարդ-թարգմանչի աշխատանքին օժանդակող միջոցներ՝ ավտոմատ բառարաններ, ստեղծելու համար, ինչպես նաև համակարգչային լեզվաբանության զարգացման համար¹: ALPAC-ի գեկույցի պատճառով ԱՄՆ-ում շուրջ մեկ տասնամյակ դադարեց մեքենական թարգմանության աշխատանքը:

Մեքենական թարգմանության անհաջողությունը ուներ մի շարք պատճառներ: Այդ տարիներին խիստ սահմանափակ էին սարքավորումների հնարավորությունները, հիշողության, արագագործության և ծրագրային միջոցները,

¹ www.altavista.com.

ինչպես նաև լեզվաբանական խնդիրների լուծման տեսական հիմունքները: Այս ամենի արդյունքում մեքենական թարգմանության համակարգը տեքստը թարգմանում էր բառացի և չէր պահպանում դրա շարահյուսական և իմաստային ամբողջությունը: Հարկ է նշել, որ մեքենական թարգմանության առաջին համակարգերը կատարում էին թարգմանություն միայն գույք լեզուների միջև /քիմար թարգմանություն/ և մշակված ձևաբանական ու շարահյուսական սահմանափակ կանոններ ունեին: Հետագայում մեքենական թարգմանության համակարգերի մշակման համար սկսեցին կիրառել կառուցվածքային և մաթեմատիկական լեզվաբանության արդյունքները: Ուշադրությունը կենտրոնացվեց ՄԹ ընդհանուր կաղապարների մշակման և կատարելագործման վրա, որպեսզի հնարավոր լինի դրանք կիրառել տարբեր լեզուների համար: Այդ ժամանակահատվածում ՄԹ զարկ ստացավ նաև ԽՍՀՄ-ում, Եվրոպայում և Ճապոնիայում: Մի քանի տարվա ընթացքում գործի դրվեց SISTRAN համակարգը USAF-ի (1970) կողմից, որը նախասահմանված է Եվրոխորհուրդի օրենսդրական և վարչական փաստաթղթերը թարգմանելու համար: Ներկայումս համակարգը աշխատում է տասնյոթ գույք լեզուներով: Թեև սկզբնական շրջանում համակարգի մեջ եղած անհաջող թարգմանության պատճառով համակարգին թերահավատորեն էին վերաբերվում, սակայն այժմ ոչ ոք չի կարող ժխտել դրա կարևորությունը և օգուտը. փաստաթղթերի արագ նախնական թարգմանությունը. պատվիրատուի մեկ հարցման թարգմանության արագությունը մոտավորապես տաս րոպե է: Այժմ SISTRAN-ը կիրառվում է երեք գլխավոր գործողությունների համար՝ ինֆորմացիային ծանոթանալու, տարբեր լեզուներով փաստաթղթեր կազմելու և թարգմանության նախնական մշակման²:

1970թ. Կանադայի Մոնրեալ համալսարանը ստեղծեց METEO համակարգը եղանակի տեսության թարգմանության համար: 1960-ական թթ. ԱՄՆ-ում և ԽՍՀՄ-ում ՄԹ հիմնականում կապված էր գիտական և տեխնիկական փաստաթղթերի ռուս-անգլերեն և անգլերեն-ռուսերեն թարգմանությունների հետ: 1970-ական թթ. կեսերին մեծացավ ՄԹ պահանջարկը՝ պայմանավորված բազմալեզու համայնքներով և բազմազգ առևտրով Եվրոպայում, Կանադայում և Ճապոնիայում:

1980-ական թթ. ՄԹ համակարգերը շատացան, շատ ավելի երկրներ ընդգրկվեցին ՄԹ ստեղծման գործում: SISTRAN-ից բացի սկսեցին ստեղծվել այլ համակարգեր ևս՝ Լոզոս համակարգը /գերմաներեն-անգլերեն և անգլերեն-ֆրանսերեն/, METAL համակարգը /գերմաներեն-անգլերեն/: Անգլերեն-ճապոներեն և ճապոներեն-անգլերեն թարգմանությունների համար ճապոնական համակարգչային կազմակերպությունները ևս ստեղծեցին համակարգեր:

² Hutchins, J. The history of machine translation in a nutshell, 2005.

1978թ-ին ՀՊՃՀ-ի պրոֆեսորադասախոսական անձնակազմի և գիտական աշխատողների միջոցով ստեղծվել է ISMA համակարգը, որի ղեկավարներն են է. Մանուկյանը և Ս. Մանուկյանը: Համակարգի նպատակն է մշակել տվյալների և գիտելիքների բազաների կառավարման և արհեստական բանականության համակարգեր: 2000թ-ից ծրագրի անձնակազմը հիմականում զբաղվել է տեքստային ինտելեկտուալ վերլուծիչների մշակմամբ:

Ղեկավարված Ազգերի Կազմակերպությանը կից համալսարանի Տոկիոյի առաջատար տեխնոլոգիաների ինստիտուտում Հ. Ուլեյիայի կողմից սկսվել են UNL միջնորդ լեզվի մշակման աշխատանքները, որին այսօր մասնակցում են Բրազիլիան, Չինաստանը, Եգիպտոսը, Ֆրանսիան, Գերմանիան, Հնդկաստանը, Ինդոնեզիան, Իտալիան, Ճապոնիան, Լատվիան, Ռուսաստանը /ընդամենը 17 երկիր/: Հայաստանը ևս անդամակցել է ընկերությանը: Հայաստանի Գիտությունների ազգային ակադեմիայի_ինֆորմատիկայի և ավտոմատացման պրոբլեմների ինստիտուտը ղեկավարում է UNL-ի գծով Հայաստանում տարվող աշխատանքները: Աշխատանքներին մասնակցում է նաև Հայաստանի Գիտությունների ազգային ակադեմիայի լեզվի ինստիտուտը:

UNL-ը Universal Networking Language՝ «Համընդհանուր ցանցային լեզու» արտահայտության հապավումն է: UNL-ը համակարգչային լեզու է, որի միջոցով հնարավոր է մշակել տեղեկատվություն և գիտելիք հաղթահարելով լեզվական պատմեշները: Այն արհեստական լեզու է, որն ընդօրինակում է բնական լեզուների հաղորդակցման գործառույթները և հնարավորություն է ընդձեռում արտահայտելու բնական լեզուների հաղորդած ամբողջ գիտելիքը: Այն համակարգչին ընձեռում է փոխհաղորդակցվելու հնարավորություն՝ մարդկանց տրամադրելով լեզվական ենթակառուցվածք՝ բազմալեզու տեղեկատվություն ստանալու, հասկանալու և հաղորդելու համար: UNL-ը բնական լեզուներից անկախ և համընդհանուր կառուցվածք է: Բնական լեզուներով արտահայտված գիտելիքը և տեղեկատվությունը փոխարկման կանոններով կարելի է արտահայտել UNL համակարգով և հակառակը, UNL-ով ներկայացված գիտելիքը և տեղեկատվությունը կարելի է վերարտադրել բնական լեզուներով: Սրա համար անհրաժեշտ նախապայման է ունենալ համակարգիչների համար մատչելի բնական լեզվի ձևային նկարագրություն և փոխարկման համապատասխան կանոններ: UNL-ն ունի բնական լեզվին համապատասխանող բոլոր բաղադրիչները: Այն կազմված է հասկացություններ արտահայտող բառերից, որոնք կոչվում են «համընդհանուր բառեր»: Համընդհանուր բառերից է կազմված UNL-ի բառարանը: Այն ոչ միայն իմաստաբանական և շարահյուսական միավոր է հասկացություն արտահայտելու համար, այլև նախադա-

³ Hiroshi Ucida, Meiying Zhu, Tarcisio G. Della Senta, Universal Networking Language, Printed in Switzerland, 2005.

սության և պարբերության UNL արտահայտությունը կառուցող հիմնական բաղադրիչ: UNL նախադասությունը կազմվում է «համընդհանուր բառերի» փոխկապակցումներով, որոնք կոչվում են «հարաբերություններ»: Իսկ հեղինակի կողմից UNL ներմուծված սուբյեկտիվ նշանակություններն արտայտվում են «բնութագրերի» միջոցով, որոնք ցույց են տալիս խոսքային գործողություն, դատողական վերաբերմունք և այլն: «Բնութագրեր»-ն օգտագործվում են նաև մի շարք հասկացական կարգեր արտահայտելու համար, օրինակ՝ սեռի քերականական կարգը:

Այսպիսով, «համընդհանուր բառեր»-ն ու «հարաբերություններ»-ը օգտագործվում են նախադասության օբյեկտիվ ինֆորմացիան նկարագրելու համար, իսկ «բնութագրեր»-ը բնորոշում են «համընդհանուր բառեր»-ը կամ իմաստային ցանցերը՝ սուբյեկտիվ տեղեկություն արտահայտելու համար:

UNL համակարգը կազմված է երեք հիմնական մասերից՝ լեզվական ռեսուրսներից, լեզվական ռեսուրսները աշխատեցնող համակարգչային ծրագրերից, լեզուն աշխատեցնող համակարգչային ծրագրերը և լեզվական ռեսուրսները պահպանող և կիրառող միջոցներից ու համակարգերից:

Լեզվական ռեսուրսները բաժանվում են լեզվից կախյալ և լեզվից անկախ մասերի: Լեզվից կախյալ ռեսուրսներ են բառային բառարանները, կանոնները և լեզվի օգտագործման համար ստեղծված համակարգչային ծրագրերը, որոնք պահվում են յուրաքանչյուր լեզվի սերվերում: Լեզվից անկախ ռեսուրս է հասկացությունների վերաբերյալ լեզվական գիտելիքը, որը պահվում է UNLKB-ի ընդհանուր հենքում: Հարկ է նշել, որ UNL տեքստի պատրաստումը ամբողջությամբ ավտոմատ չէ, ամբողջ աշխատանքը բաժանված է համակարգչի և մարդ-մասնագետի միջև: Սրանով UNL-ը տարբերվում է մեքենական մյուս թարգմանություններից: Ինչպես տեսնում ենք, ՄԹ համար ստեղծված յուրաքանչյուր համակարգ կամ դեռևս ծրագիր չի կատարում կատարյալ կամ զոնե ճշգրիտ թարգմանություն: Իհարկե, ՄԹ յուրաքանչյուր ծրագրի անկատար թարգմանությունը պայմանավորված է տարբեր պատճառներով կամ խնդիրներով: Մեր նշած ծրագրերի մեջ կարող ենք առանձնացնել մի բանի պատճառներ, բայց նախ նշենք, որ ցանկացած ՄԹ ծրագրի համար անհրաժեշտ է կոնկրետ լեզվի ձևային նկարագրություն և փոխարկման համապատասխան կանոններ: Հայերենի քերականության ձևային կամ կաղապարային նկարագրության ուղղությամբ նախկինում կատարվել են որոշակի աշխատանքներ⁴. Գ.

⁴ Г. Джаукян, Универсальная теория языка, М., 1999, *Մ. Գալստյան, Ժամանակակից հայերենի բաղադրյալ բառերի և նրանց բանաձևերի մասին*, ԲԵՂ, 1972, հ.3, *Մ. Գալստյան*, Ածանցումը և ածանցները ժամանակակից հայերենում, Ե. 1978, *Ռ. Ոստույան*, Ժամանակակից հայերենի անվանական բառափոխման ձևային նկարագրություն, «Լեզվի և ոճի հարյեր», ժող., պ. VI, Ե., 1982, *Ռ. Ոստույան*, Ժամանակակից հայերենի ածանցման ձևային նկարագրություն, «Բնո, նախա-

Զահուկյանը նշված աշխատությունում նշակել է համընդհանուր լեզվաբանական կադապարը և մասամբ կիրառել հայերենի նկատմամբ: Սակայն որևէ լեզվի քերականության ձևային ամբողջական նկարագրություն ունենալու համար պարտադիր է քերականական կառուցվածքի բոլոր մակարդակների ներառումը տվյալ լեզվով տեղեկատվության ստացման, պահպանման, մշակման, փոխակերպման և հաղորդման պահանջները առավելագույնս բավարարելու համար:

ՄԹ ծրագրերի նկարագրությունից պարզ է առաջացած խնդիրների բնույթը, որը հիմնականում պայմանավորված է քերականության ձևային նկարագրության խնդիրներով: Ինչպես գիտենք, յուրաքանչյուր բնական լեզու բարդ և ոչ միատարր կառուցվածք ունի և ձևայնացումը այնքան էլ հեշտ չէ: Բնական լեզվի ձևայնացումը զգալի բարդություններ ունի ամենից առաջ ձևերի բազմիմաստության և միևնույն իմաստների բազմաձևության պատճառով: Այդ դժվարությունները հաղթահարելու համար պետք է ընտրել որոշակի սկզբունքներ: Նախ, նկարագրության օբյեկտը պետք է լինի կանոնարկված զրական լեզուն առանց բարբառային, ժարգոնային, հնացած ձևերի, այլ կերպ ասած լեզվի ստանդարտ վիճակը: Ապա՝ նկարագրությունը պետք է լինի լրիվ, ներառելով բոլոր բացառությունները, անկանոն ձևերն ու շեղումները (նկատի ունենք միևնույն համակարգի սահմաններում), նկարագրությունը պետք է լինի անհակասական և իզոմորֆ, այսինքն լեզվի տարբեր մակարդակներ պետք է ներկայացվեն միևնույն միասնական սկզբունքներով և վերջապես, նկարագրությունը պետք է լինի խնայողական, այն իմաստով, որ սահմանափակ թվով կանոնների օգնությամբ հնարավոր լինի ներկայացնել միավորների առավելագույն քանակ⁵: Նշենք ձևային նկարագրության փորձերի մեջ առաջացած խնդիրներ. ա/ Մինչ այս արված փորձերն ունեն մասնավոր խնդիրներ կամ հատվածական են. բ/ Նախատեսված չեն համակարգչային ծրագրերին ծառայելու համար. գ/ Նկարագրությունները լրիվ կամ ամբողջական չեն. դ/ Ներառված չեն բոլոր բացառությունները, անկանոն ձևերն ու շեղումները. ե/ Նկարագրությունները խնայողական չեն, այսինքն՝ միավորների առավելագույն քանակը ներկայացված չէ սահմանափակ թվով կանոնների օգնությամբ:

Սակայն, ինչպես նկարագրությունն է ցույց տալիս, ՄԹ նշված ծրագրերից և ոչ մեկը դեռևս չի կատարում ճշգրիտ թարգմանություն և կարծես յուրաքանչյուրում առկա են հատկապես շարահյուսական խնդիրները, ինչպես նաև իմաստաբանական խնդիրները:

դասություն, Կոնստ, ժող. Ե. 1989, *Ժ. Մանուկյան «Խոսել»* բայի իմաստաբանական հարկանիշները, «Լեզվի և ոճի հարցեր», ժող., պ. VI, Ե., 1982, *Լ. Նովսեփյան*, Գրաբարի բառակազմությունը, Ե., 1987, *A. Manukyan, A Formalised Model for Ancient Armenian Verb*, Proceedings of the International Conference on Computer Science and Information Technologies, Erevan, 2003:

⁵ Լ. Նովսեփյան, Արևելագիտության հարցեր, VI, Եր., 2006, էջ 146: