

## ՊԱՇՏՈՆԱԿԱՆ ԸՆԴԴԻՄԱԽՈՍԻ ԿԱՐԾԻՔԸ

Կարեն Հրահատի Նիկողոսյանի «Հայոց լեզվով գրավոր տեքստը բանավոր խոսքի ձևափոխման ավտոմատացված համակարգի մշակումը» թեմայով  
Ե.13.02 - «Ավտոմատացման համակարգեր» մասնագիտությամբ տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի հայցման ատենախոսության վերաբերյալ

Կարեն Նիկողոսյանի «Հայոց լեզվով գրավոր տեքստը բանավոր խոսքի ձևափոխման ավտոմատացված համակարգի մշակումը» թեմայով ատենախոսությունը կազմված է ներածությունից, 4 գլուխներից, եզրահանգումից, 110 աղբյուր ընդգրկող գրականության ցանկից, 2 հավելվածներից: Հետազոտության հիմնական տեքստը շարադրված է 145 էջում:

**Թեմայի արդիականությունը** պայմանավորված է տեքստից խոսքի վերափոխման դժվար և ոչ լիովին լուծված խնդրի ուսումնասիրման անհրաժեշտությամբ: Արհեստական բանականության (ԱԲ) ալգորիթմների զարգացումը հնարավորություն է տվել բարելավել խոսքի սինթեզի ճշգրտությունը և կիրառել դրանք հայոց լեզվով տեքստը խոսքի վերափոխող համակարգերում: Ատենախոսական աշխատանքը նվիրված է հայոց լեզվով տեքստը խոսքի վերափոխող համակարգի մշակմանը՝ հաշվի առնելով հայերենի հնչյունաբանական առանձնահատկությունները: Օպտիկական նիշերի ճանաչման մոդելի ինտեգրումը հնարավորություն է տվել համակարգին աշխատել նաև սկանավորված փաստաթղթերի և լուսանկարված նյութերի հետ՝ ընդլայնելով համակարգի կիրառման հնարավորությունները կրթական, հանրային ծառայությունների և այլ ոլորտներում: Ստեղծված բոլոր բաղադրիչները կխթանեն հայոց լեզվի համար խոսքի սինթեզի տեխնոլոգիաների հետագա զարգացումը:

### **Աշխատանքի համառոտ բովանդակությունը**

Ներածությունում հիմնավորված է ատենախոսության թեմայի արդիականությունը, ներկայացված են ատենախոսության նպատակները, պաշտպանության դուրս բերվող հիմնական դրույթները, աշխատանքի հակիրճ նկարագրությունը և հիմնական գիտական արդյունքները:

Առաջին գլխում ներկայացված են գրավոր տեքստը բանավոր խոսքի վերափոխման ավտոմատացված համակարգերի ընդհանուր դրույթները: Նախ հիմնավորված է գրավոր տեքստը բանավոր խոսքի վերափոխման ավտոմատացված համակարգի մշակման անհրաժեշտությունը, քննարկված են նմանատիպ

ավտոմատացված համակարգերի առկա միջոցները՝ խորը ուսուցման ճարտարապետությունները, դրանց համեմատական վերլուծությունը, գրանշանները հնչյունների վերափոխող համակարգերի արդյունավետության գնահատման մեթոդները, այնուհետև առանձնացված են հայոց լեզվով գրավոր տեքստը բանավոր խոսքի վերափոխման ավտոմատացված համակարգերում առկա խնդիրները: Կատարված վերլուծության հիման վրա ձևակերպված են ատենախոսության նպատակները:

Երկրորդ գլուխը նվիրված է ավտոմատացված համակարգի ստեղծման համար անհրաժեշտ տվյալների հավաքածուների հավաքագրմանը, մշակմանը և քանակական բնութագրերի վերլուծությանը: Ներկայացված է ձայնային, ինչպես նաև գրանշանները հնչյունների վերափոխող տվյալների հավաքածուների հավաքագրման, մշակման ու վերլուծության գործընթացները:

Երրորդ գլխում նկարագրված են հայոց լեզվով գրավոր տեքստը բանավոր խոսքի վերափոխման և գրանշանները հնչյունների վերափոխման մաթեմատիկական մոդելները, ուսուցման գործընթացը: Նկարագրված է նաև տեքստը խոսքի վերափոխող (ՏԽՎ-ՏՏՏ) ԱԲ մոդելի օպտիմալացումն ու քվանտացումը, ամփոփված են նշված գործընթացների արդյունքները: Այս գլխում ներկայացված է ավտոմատացված համակարգի մշակված ծրագրային միջոցը, դրա հիմնական բաղադրիչները, տրված են համակարգն օգտագործելու քայլերը: Նկարագրված է նաև համակարգի «Image to Text» ներդիրը, որը նախատեսված է գրավոր տեքստ պարունակող պատկերներում առկա հայերեն տեքստերը մշակելու, ընթերցելու համար:

Չորրորդ գլխում ներկայացված են փորձարարական հետազոտությունների արդյունքները: Վերլուծված են հայոց լեզվով բազմախոսնակային հակամարտ ուսուցմամբ ամբողջական վարիացիոն ավտոկոդավորիչների ճարտարապետությամբ տեքստից խոսքի վերափոխման ԱԲ մոդելի գնահատման արդյունքները՝ խոսնակի նույնականության պահպանման, ընկալման որակի, հասկանալիության գնահատումները, ինտոնացիոն և հնչյունային ճշգրտության գնահատումը, չափանիշների համադրված վերլուծությունը և խոսնակների համեմատությունը: Ներկայացված է նաև օպտիմալացված բազմախոսնակային ՏԽՎ ԱԲ մոդելի գնահատման արդյունքների վերլուծությունը, ընդ որում ԱԲ և ԱԲ ՏԽՎ մոդելների կատարողականի վերլուծությունը բազմամիջուկ ապարատային կարգավորման պայմաններում:

Եզրահանգման և առաջարկությունների մեջ ընդհանրացված են աշխատանքում ստացված գիտական ու գործնական արդյունքները և տրված են կիրառման երաշխավորությունները:

Հավելված 1-ում Հայաստանի ազգային պոլիտեխնիկական համալսարանի «Տեղեկատվական տեխնոլոգիաներ և ավտոմատացում» ամբիոնում դասավանդվող «Բնական լեզվի մշակում», «Բնական լեզվի բանական վերլուծության մեթոդներ» դասընթացներում ատենախոսական հետազոտության արդյունքների կիրառումը հավաստող ներդրման ակտն է:

Հավելված 2-ում տրված է հակամարտ ուսուցմամբ ամբողջական վարիացիոն ավտոկոդավորիչների ճարտարապետությամբ ՏԽՎ մոդելի պարամետրերի նկարագրությունը:

#### **Ատենախոսության գիտական նորույթ:**

- Մշակվել է Conformer-CTC ճարտարապետությամբ գրանշանները հնչյունների վերափոխող մոդել հայոց լեզվի համար, որն ապահովում է 16.13% բառային սխալի մակարդակ և 17.36% հնչյունային սխալի մակարդակ, զգալիորեն գերազանցելով առկա Phonemizer գործիքի արդյունքները:
- Մշակվել է ձայնային տվյալների հավաքագրման ավտոմատացված համակարգ, որն իրագործում է dBFS սանդղակով ազդանշանի մակարդակի հաշվարկ և դինամիկ շեմային մեթոդ՝ օպտիմալ բաժանման կետերը հայտնաբերելու համար: Համակարգը զգալիորեն արագացնում է տվյալների հավաքագրումը՝ առանց ֆինանսական ռեսուրսների պահանջի:
- Ստեղծվել են ձայնային տվյալների հավաքածու (ընդհանուր 21 ժամ ձայնագրություն կամ 14,182 ձայնային ֆայլ) և գրանշանները հնչյունների վերափոխող տվյալների հավաքածու (17,862 գրանշան-հնչյուն զույգ):
- Մշակվել է բազմախոսնակային հակամարտ ուսուցմամբ ամբողջական վարիացիոն ավտոկոդավորիչների ճարտարապետությամբ տեքստից խոսքի վերափոխման մոդել հայոց լեզվի համար, որը ցուցաբերում է հետևյալ արդյունքները՝ խոսնակի նույնականության գործակցի 0.9366 միջին արժեք, F0-կոռելյացիայի 0.8834 արժեք, խոսքի ընկալողական գնահատականի 2.87 միջին արժեք և կարճաժամկետ օբյեկտիվ ըմբռնելիության 0.7248 արժեք:
- ONNX ձևաչափի և քվանտացման միջոցով տեքստից խոսքի վերափոխման մոդելի օպտիմալացումը հանգեցրել է. մոդելի ծավալի 87%-ով նվազեցման (1.0 ԳԲ-ից մինչև 127.5 ՄԲ), արձագանքի ժամանակի մոտավորապես 85%-ով

շի  
խ  
ած  
իկ

կրճատման և իրական ժամանակի գործոնի 0.22-0.30 միջակայքում արժեքների ապահովման:

- Մշակվել է օպտիկական նիշերի ճանաչման Tesseract համակարգի հետ ինտեգրման նոր մեթոդ, որում օգտագործվել է ավտոմատ գեներացված սինթետիկ ուսուցման տվյալներով հարստացված մոդել: Ձևավորված մոդելն ուսուցանվել է բարդ և բազմազան օրինակների վրա, որոնք պարունակում են տարբեր տեսակի աղմուկներ, լրացուցիչ տառատեսակներ և տեքստի բարձրության տարբերակներ: Հայալեզու տեքստերի վրա առաջարկված մոտեցումը թույլ է տվել կրճատել ճանաչման սխալների գործակիցը 56%-ով՝ ինչպես բառային, այնպես էլ նիշային մակարդակներում: Համակարգն ապահովում է բարձր ճշգրտություն աղմկոտ ձայնագրությունների և ցածրորակ պատկերների դեպքում՝ էականորեն ընդլայնելով տեքստը խոսքի վերափոխող համակարգի կիրառելիությունը:

**Գիտական դրույթների, եզրակացությունների հիմնավորման և հավաստության աստիճանը հաստատված են** ատենախոսությունում ներկայացված գրանշանները հնչյունների վերափոխող և տեքստը խոսքի վերափոխող արհեստական բանականության մոդելների ծրագրային իրագործման, համակողմանի փորձարարական հետազոտությունների և միջազգայնորեն ընդունված չափանիշների միջոցով ստացված գնահատումների արդյունքներով և մաթեմատիկական հիմնավորումներով:

**Ատենախոսության հիմնական դրույթների արտացոլման աստիճանը սեղմագրում**

Աշխատանքի հիմնական գիտական և գործնական արդյունքները լիովին արտացոլված են սեղմագրում:

**Աշխատանքի գործնական նշանակությունը**

Ատենախոսությունում մշակված հայոց լեզվով տեքստը խոսքի վերափոխող համակարգն իրականացվել է ինչպես սերվերային, այնպես էլ ներդրված սարքերում կիրառելու նպատակով: Համակարգն ապահովում է բարձր որակի խոսքի սինթեզ: Ներկառուցված սարքերի համար ONNX ձևաչափի և քվանտացման միջոցով կատարված օպտիմալացումը հնարավորություն է տվել նվազեցնել մոդելի ծավալը, կրճատել արձագանքի ժամանակը և աշխատել իրական ժամանակից 3-4 անգամ արագ: Օպտիկական նիշերի ճանաչման բաղադրիչի ինտեգրումը, որը սինթետիկ ուսուցման տվյալների շնորհիվ կրճատել է ճանաչման սխալների գործակից արժեքը 56%-ով, զգալիորեն կարող է ընդլայնել համակարգի կիրառելիությունը: React/Flask տեխնոլոգիաներով օգտագործողի գրաֆիկական միջավայրը համակարգը դարձրել

է հասանելի ինչպես մասնագետների, այնպես էլ առանց փորձառության օգտագործողների համար: Մշակված համակարգը կարող է կիրառվել կրթական նյութերի ձայնային տարբերակների ստեղծման, գրադարանային համակարգերում ձայնային գրքերի նախապատրաստման, լրատվական ոլորտում ձայնային բովանդակության ստեղծման և հաճախորդների սպասարկման ոլորտում ԱԲ գործակալների մշակման համար:

### **Ընդհանուր գնահատականներ**

Ատենախոսության հիմնական արդյունքները հրատարակված են 9 գիտական աշխատանքներում, որոնք ներկայացված են գրականության ցանկում:

### **Դիտողություններ և առաջարկություններ**

- աշխատանքում նշվում է, որ ստեղծվել է ազատ հասանելիությամբ կայքէջ, սակայն հղումը տրված չէ,
- ցանկալի կլիներ, որ լիներ համեմատական վերլուծություն այսօր առկա ազատ հասանելիությամբ հետևյալ համակարգերի հետ՝ <https://www.naturalreaders.com/online/>, <https://elevenlabs.io/text-to-speech> <https://www.narakeet.com/languages/armenian-text-to-speech/>,
- հստակ ներկայացված չեն ըստ տեքստի բովանդակության բանավոր խոսքի առոգանության, մեղեդայնության, հնչերանգի, շեշտադրումների որոշման առանձնահատկությունները,
- «բազմախոսնակային տեքստը խոսքի վերափոխման համակարգ» արտահայտության փոխարեն ճիշտ կլիներ «տեքստը խոսքի վերափոխման բազմախոսնակային համակարգ» - ի օգտագործումը, «մոդելի չափի» փոխարեն «մոդելի ծավալ» բառակապակցության օգտագործումը, «հակառակորդային ուսուցման» փոխարեն՝ «հակամարտ ուսուցում», «ընկալելիության» փոխարեն՝ «ըմբռնելիություն» և այլն: Օգտագործվում են և՛ «դիկտոր», և՛ «խոսնակ» բառերը,
- ցանկալի է ընդլայնել տվյալների հավաքածուն ներառելով իգական, մանկական ձայներ և բարբառային տարբերակներ,
- աշխատանքում ներկայացված են հիմնականում օբյեկտիվ չափանիշներ, սակայն բացակայում է սուբյեկտիվ գնահատումը (օրինակ՝ օգտագործողների հարցումներ) սինթեզված խոսքի բնականության վերաբերյալ: Ցանկալի է անցկացնել լրացուցիչ հետազոտություններ՝ ներառելով MOS (Mean Opinion Score) գնահատականներ:

**Եզրակացություն:** Կարեն Հրահատի Նիկողոսյանի «Հայոց լեզվով գրավոր տեքստը բանավոր խոսքի ձևափոխման ավտոմատացված համակարգի մշակումը» թեմայով ատենախոսությունը արդիական է, կատարված է բարձր գիտական մակարդակով, ստացված արդյունքները հավաստի են, ունեն տեսական և կիրառական արժեք: Նշված դիտողությունները չեն նսեմացնում աշխատանքի արժանիքները, ուստի կարծում եմ, որ ատենախոսությունը բավարարում է ՀՀ-ում գիտական աստիճանների շնորհման կարգի թեկնածուական ատենախոսություններին ներկայացվող բոլոր պահանջներին, իսկ հեղինակը՝ Կարեն Հրահատի Նիկողոսյանն արժանի է տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի շնորհմանը Ե.13.02 – «Ավտոմատացման համակարգեր» մասնագիտությամբ:

Պաշտոնական ընդդիմախոս,  
տ.գ.դ., պրոֆ.

*Ս. Սարգսյան*  
9.07.25թ.

Ա.Գ. Ավետիսյան

տ.գ.դ., պրոֆ. Ա.Գ. Ավետիսյանի ստորագրությունը հաստատում եմ  
ՀԱՊՀ գիտական քարտուղար Ծ. Հովհաննիսյան

