Дүйсенбекқызы Жансаяның 8D06101 – Ақпараттық жүйелер білім беру бағдарламасы бойынша философия докторы (PhD) дәрежесін алу үшін ұсынылған «Сөйлеу қабілеті бұзылған балалардың қазақша сөйлеуді тану әдісін зерттеу және әзірлеу» тақырыбындағы диссертациялық жұмысына

АҢДАТПА

Зерттеу тақырыбының өзектілігі.

Ғылыми-технологиялық дамудың қазіргі кезеңінде сөйлеуді автоматты тану (Automatic Speech Recognition, ASR) жүйелері білім беру платформаларында, логопедиялық диагностика құралдарында және ерте балалық шақты дамытуға бағытталған ақпараттық жүйелерде кеңінен қолданыс тапқан. Дегенмен, сөйлеу қабілеті бұзылған балалардың сөйлеуін автоматты тану мәселесі ғылыми тұрғыдан жеткіліксіз зерттелген және практикалық шешімі толық табылмаған күрделі ғылыми-техникалық проблема болып қала береді.

Балалар мен ересектер сөйлеуі арасындағы акустикалық айырмашылықтар эмпирикалық зерттеулермен дәлелденген. Атап айтқанда, балалардың негізгі жиілігі (F0) 250–400 Гц шегінде, ал ересектерде бұл көрсеткіш 100–150 Гц деңгейінде болады; форманттық жиіліктер де айтарлықтай жоғары — балалардың дауыс жолы ұзындығы 9–13 см болса, ересектерде бұл 17 см-ге жетеді. Толық артикуляция 12–14 жасқа жеткенде ғана қалыптасады. Зерттеулер дәлелдегендей, ересек сөйлеуіне негізделген ASR модельдері балалар сөйлеуіне қолданылғанда сөздердегі қате мөлшері (Word Error Rate, WER) 40–45%-ға өсетіні анықталған.

Қазақ тілі — агглютинативті морфологиялық құрылымға ие, ресурстары шектеулі (low-resource) тіл. Қолданыстағы қазақстандық ASR жүйелері (KATEK, KazASR, KazakhTTS, KazakhBERT, Soyle) тек ересектер сөйлеуіне негізделген; балалар сөйлеуіне арналған мамандандырылған аннотацияланған акустикалық корпус 2025 жылға дейін болмаған. Бұдан басқа, Kaldi, Mozilla DeepSpeech және Google Speech-to-Text жүйелерінің тәуелсіз бағалауы бойынша ең жақсы нәтижені Google STT WER = 52,97% деңгейінде ғана көрсеткен, бұл практикалық қолданыс үшін жеткіліксіз болып табылады.

Жоғарыда баяндалған ғылыми дәлелдемелер мен қолданыстағы жүйелердің шектеулері сөйлеу қабілеті бұзылған балалардың қазақша сөйлеуін автоматты тану мен ерте диагностикалауға арналған ғылыми-техникалық шешімнің өзектілігін айқындайды. Ұсынылып отырған диссертациялық жұмыста қазақ тіліндегі балалар сөйлеуіне арналған аннотацияланған акустикалық датасет алғаш рет жасалып, ASR мен LLM

интеграциясына негізделген комплексті post-editing тәсілі ұсынылады. Жүргізілген эксперименттер ұсынылған тәсілдің жоғары дәлдікке, есептеу тиімділігіне және практикалық қолдану мүмкіндіктеріне ие екенін дәлелдейді.

Диссертациялық жұмыстың мақсаты

Қазақ тіліндегі балалар сөйлеуіне арналған аннотацияланған акустикалық деректер қорын тұңғыш рет құру және сөйлеу қабілеті бұзылған балалардың сөйлеуін тану дәлдігін арттыруға бағытталған ASR мен LLM интеграциясына негізделген комплексті post-editing әдісін әзірлеу.

Зерттеу міндеттері:

–дүниежүзіндегі балалар акустикалық аудиодеректеріне салыстырмалы-аналитикалық шолу жүргізу және олардың құрылымдық, көлемдік, акустикалық ерекшеліктерін жүйелеу; балалар сөйлеуін тану саласындағы заманауи әдістер мен акустикалық корпустарға салыстырмалы талдау жүргізу;

–балалардың жас ерекшелігі мен жыныстық сипаттамаларын ескере отырып, қазақ тіліндегі балалар сөйлеуіне арналған аннотацияланған акустикалық датасет жинақтау және өңдеу;

–заманауи нейрондық ASR модельдерін (Whisper, MMS, TurkicASR, ESPnet, DeepSpeech және т.б.) зерттеу және оларды балалар сөйлеуіне fine-tuning арқылы бейімдеу;

–қосымша оқытудан кейінгі ASR нәтижелерін түзетуге арналған LLM негізіндегі комплексті post-editing тәсілін әзірлеу;

–ұсынылған әдістің тиімділігін эксперименттік тұрғыдан бағалау және салыстырмалы талдау жүргізу.

Зерттеу әдістері:

Машиналық оқыту және терең оқыту әдістері, сөйлеуді автоматты тану (ASR) технологиялары, акустикалық сигналдарды өңдеу, ықтималдықтар теориясы мен математикалық статистика, бағдарламалық жасақтаманы әзірлеу әдістері.

Зерттеу нысаны

Қазақ тіліндегі сөйлеу қабілеті бұзылған балалардың сөйлеу үдерісі және оның акустикалық ерекшеліктері.

Зерттеу пәні

Қазақ тіліндегі сөйлеу қабілеті бұзылған балалардың сөйлеуін тануға арналған технологиялар, алгоритмдер, модельдер мен әдістер, бағдарламалық жасақтама.

Алынған нәтижелердің ғылыми жаңалығы:

1. Қазақ тілінде сөйлейтін балалардың жас, жыныс және фонетикалық ерекшеліктерін ескеретін арнайы аннотацияланған акустикалық датасет алғаш рет әзірленді. Жасалған деректер қоры балалар сөйлеуін зерттеу саласындағы ашық қолжетімді ресурстардың тапшылығын жоюға маңызды үлес қосады.

2. Сөйлеу қабілеті бұзылған балалардың қазақша сөйлеуін тануға арналған, ASR моделін fine-tuning арқылы бейімдеуді және LLM негізіндегі постөңдеуді біріктіретін комплексті post-editing тәсілі ұсынылды. Аталған тәсіл балалардың қазақша сөйлеуді тану дәлдігін арттырады.

Қорғауға шығарылатын тұжырымдар:

1. Жасалған аннотацияланған акустикалық датасет балалардың жас, жыныс және фонетикалық ерекшеліктерін ескере отырып жинақталған. Бұл датасет қазақ тіліндегі балалар сөйлеуін зерттеуге арналған алғашқы мамандандырылған ресурс болып табылады.

2. Ұсынылған комплексті post-editing тәсілі ASR моделін fine-tuning арқылы балалар сөйлеуіне бейімдеуді және LLM негізіндегі постөңдеуді біріктіреді. Жүргізілген эксперименттер нәтижесінде WER 0,70-тен 0,18-ге, CER 0,467-ден 0,065-ке төмендетілді.

Диссертациялық жұмыстың көлемі мен құрылымы

Диссертация кіріспеден, төрт негізгі тараудан, қорытындыдан, пайдаланылған әдебиеттер тізімінен және қосымшалардан тұрады. Жұмыстың жалпы көлемі — 123 бет машиналық мәтін, оның ішінде 30 сурет, 43 кесте, 117 пайдаланылған әдебиет және 4 қосымша.

Бірінші тарауда сөйлеуді тану және сөйлеу бұзылыстары саласындағы зерттеулерге аналитикалық шолу жасалады: ASR жүйелерінің даму эволюциясы, балалар сөйлеуінің акустикалық-фонетикалық ерекшеліктері, сөйлеу бұзылыстарының жіктелуі және аз ресурсты тілдердегі халықаралық тәжірибе жүйелі түрде қарастырылады.

Екінші тарауда мәтіндік материалды іріктеу әдіснамасы, акустикалық деректерді жинау мен аннотациялау әдістері, этикалық және құқықтық аспектілер баяндалады; датасет құрылымы мен деректерді алдын ала өңдеу әдістері ұсынылады. Аудиодеректер ҚР «Дербес деректер және оларды қорғау туралы» Заңына (№ 94-V) сәйкес ата-аналардың немесе заңды өкілдердің жазбаша келісімі негізінде жиналды; барлық жазбалар бірегей идентификаторлар беру арқылы депсонализацияланып, Қазақстан Республикасының аумағындағы қорғалған серверлерде шектеулі қолжетімділікпен сақталады. Қатысушылардың дербес деректері жарияланбайды және үшінші тұлғаларға берілмейді.

Үшінші тарауда ұсынылған комплексті әдістің архитектуралық құрылымы, сөйлеуді тану есебінің формализациясы, балалар сөйлеуіне бейімделген акустикалық модельдің математикалық қойылымы және комплексті әдістің математикалық моделі баяндалады.

Төртінші тарауда ұсынылған әдістің эксперименттік бағалауы жүргізіледі: ASR модельдерінің салыстырмалы талдауы, fine-tuning және LLM негізінде жасалған post-editing нәтижелері, сондай-ақ ұсынылған әдістің практикалық ортада апробациясының нәтижелері ұсынылады.

Жұмыстың апробациясы

Зерттеу нәтижелері әл-Фараби атындағы Қазақ ұлттық университетінің ғылыми семинарларында талқыланып, халықаралық конференцияларда баяндалды, сондай-ақ №154 балабақшада тәжірибелік апробациядан өтті.

Жарияланымдар.

Диссертация тақырыбы бойынша 9 ғылыми жұмыс жарияланды, оның ішінде 2 жұмыс Web of Science (SCIE, Q2) және (ESCI, Q1) базаларына индексделетін журналдарда, 5 жұмыс халықаралық конференция жинағында, 2 жұмыс ҚР БҒМ БҒСБК ұсынатын журналдарда жарияланды; 2 авторлық куәлік алынды.

- Rakhimova D., Duisenbekkyzy Zh., Adali E. Investigation of ASR Models for Low-Resource Kazakh Child Speech: Corpus Development, Model Adaptation, and Evaluation// Applied Sciences. – 2025. – Vol. 15, Issue 16, Article 8989. DOI: 10.3390/15168989 (WoS SCIE, Q2, IF: 2,7, процентиль 79).
- Rakhimova D., Karibayeva A., Turarbek A., Karyukin V.I., Duisenbekkyzy Zh. and Aliyev R. Development of a Children's Educational Dictionary for a Low-Resource Language Using AI Tools // Computers. – 2024. – Vol. 13, No. 10, Article 253. DOI: 10.3390/computers13100253 (WoS ESCI, Q1, CiteScore: 7,5, процентиль 84).
- Rakhimova, D. Duisenbekkyzy Zh., Research on the Impact of Social Networks on the Development of Bilingualism in Kazakhstan. Advances in Intelligent Systems and Computing, 2024, pp. 123–134. Percentile: 39 (Springer)
- Rakhimova, D., Yessirkepova, A., Adali, E. Fine-Tuning Neural ASR Models for Low-Resource Kazakh Children’s Speech with Preprocessing Enhancements. 10th International Conference on Computer Science and Engineering (UBMK’2025), Istanbul, Türkiye, 17–19 September 2025, pp. 428–432. (Scopus Proceeding)
- Рахимова, Д. Р., Дүйсенбекқызы Ж., Шадияр, Т. Қазақ тіліндегі балалар корпусы үшін деректерді жинау және өңдеу. Студенттер мен жас ғалымдардың «ФАРАБИ ӘЛЕМІ» халықаралық ғылыми конференциясы, Алматы, 4–6 сәуір 2024 ж.
- Рахимова, Д. Р., Дүйсенбекқызы Ж., Турганбаева, М. Ж., Арабов, Б. Б. Балалар акустикалық корпусын қазақ тілінде жинау және өңдеу жұмысы. «ǴYLYM JÁNE BILIM – 2024» халықаралық ғылыми конференциясы, сәуір 2024 ж.
- Дүйсенбекқызы Ж., Рахимова Д. Р., Есиркепова А. П. Төмен ресурсты қазақ тілінде балалар сөйлеуін тануға арналған нейрондық ASR үлгілерін бейімдеу. «Информатика және қолданбалы математика» Х

Халықаралық ғылыми-практикалық конференция материалдары, Алматы, Қазақстан, 8–11 қазан 2025 ж., Б. 303–309.

– Д.Р. Рахимова, Ж. Дүйсенбекқызы, О.А. Сағынтай, Ә. Турарбек, Ш. Төлеуғали. Балалардың сөйлеуін тану жүйелеріндегі шуды азайту алгоритмдерін талдау // ҚазҰПУ хабаршысы. – 2025. – №91(3). – Б. 230–242. DOI: 10.51889/2959-5894.2025.91.3.021

– Rakhimova D., Duisenbekkyzy Zh., Adali E., Toleugali Sh. Research and Processing of the Kazakh Children's Acoustic Corpus // ҚР ҰИА Хабаршысы. – 2024. – №3(93).

– Авторлық куәлік «Dataset Loader Bot». Авторлық құқық объектісі №56775. Тіркелген: 15 сәуір 2025 ж. Авторлар: Рахимова Д., Дүйсенбекқызы Ж., Алиев Р.

– Авторлық куәлік «Қазақ тіліндегі 2–9 жас аралығындағы балалар сөйлеу аудиожазбаларының таңбаланған акустикалық деректер базасы». Авторлық құқық объектісі №70955. Тіркелген: 27 сәуір 2026 ж. Авторлар: Дүйсенбекқызы Ж., Рахимова Д.