ISNN 3030-3907

2026
MAY
ISSUE - 5
VOLUME - 2

DEVELOPMENT OF SCIENCE

scientific journal



Development of science

Ilmiy jurnal

MAY 2026
Issue - 5, VOLUME - 2

ISSN 3030 -3907

Ilmiy jurnal OAK rayosatining 2024 yil 27 sentabrdagi 361-son qaroriga asosan 02.00.00 kimyo fanlari hamda 2025 yil 12 fevral 367- son qaroriga asosan 05.00.00 texnika fanlari bo‘yicha falsafa doktori (PhD) va fan doktori (DSc) ilmiy darajasiga talabgorlarning dissertatsiya ishlari yuzasidan dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan milliy ilmiy nashrlar ro‘yxatiga kiritilgan.

BOSH MUHARRIR:
Temirov Alisher Hoshim o‘g‘li – texnika fanlari falsafa doktori (PhD),Dotsent.
TAHRIRIYAT KENGASH RAISI:
Fozilov Sadriddin Fayzullayevich – texnika fanlari doktori, professor
TAHRIRIYAT KENGASH A’ZOLARI

TEXNIKA FANLARI/ TECHNICAL SCIENCES/ ТЕХНИЧЕСКИЕ НАУКИ

Bozorov G‘ayrat Rashidovich
Buxoro Davlat Texnika Universiteti, professor
Abdurahmonov Olim Rustamovich
Buxoro Davlat Texnika Universiteti, professor
Adizov Bobur Zamirovich
O‘zbekiston Respublikasi Fanlar akademiyasi Umumiy va noorganik kimyo instituti, professor
Hayitov Ruslan Rustamjonovich
Buxoro Davlat Texnika Universiteti, professor
Jumayev Qayum Karimovich
Buxoro Davlat Texnika Universiteti, professor.
Berdiyev Dorob Murodovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor
Maxmudov Rafiq Amonovich
Buxoro Davlat Texnika Universiteti, professor.
Tilloyev Lochin Ismatilloyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Raxmonov Qahramon Sanakulovich
Buxoro Davlat Texnika Universiteti, dotsent.
Barakayev Nusratullo Rajabovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor.
Jo‘rayev Xayrullo Fayziyevich
Buxoro Davlat Texnika Universiteti, professor.
Panoyev Erali Rajabboyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Ataullayev Sherzod Nabiullyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Malik Negmurodovich
Buxoro Davlat Texnika Universiteti, dotsent.
Axmedova Ozoda Bahronovna
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti.
Safarov Bahri Jumayevich
Buxoro Davlat Texnika Universiteti, dotsent.
Xaydarov Axtam Amonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Safarov Jasur Alijon o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.

KIMYO FANLARI/ CHEMICAL SCIENCES/ ХИМИЧЕСКИЕ НАУКИ

Axmedov Voxid Nizomovich
Buxoro Davlat Texnika Universiteti, professor
Umarov Baqo Bafoyevich
Buxoro Davlat Universiteti, professor
Kodirov Abduaxad Abduraximovich
Qarshi davlat Universiteti, professor
Qurbonov Mingniqul Jumagulovich
Qarshi davlat Universiteti, dotsent
Turemuratov Sharibay Naurizbaevich
O‘zbekiston Respublikasi Fanlar akademiyasi Qoraqalpog‘iston bo‘limi Bosh ilmiy kotibi, professor
Tursunov Kaxor Shonazarovich
Qarshi davlat Universiteti, professor
Do‘stov Hamro
Buxoro Davlat Texnika Universiteti, professor
Panoyev Nodir Shavkatovich
Buxoro Davlat Texnika Universiteti, dotsent.
Olimov Bobur Bahodir o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.
Kasimov Sherzod Abduazimovich
Termiz Davlat Universiteti, professor.
Aliqulov Rustam Valiyevich
Termiz Davlat Universiteti, professor
Tropov Xamza Tursunovich
Samarqand Davlat Universiteti, professor

IQTISODIYOT FANLARI/ ECONOMICS/ ЭКОНОМИЧЕСКИЕ НАУКИ

Nizamov Aslitdin Badreddinovich
Buxoro Davlat Texnika Universiteti, professor.
Boboyev Akmal Choriyevich
Buxoro Davlat Texnika Universiteti, professor.
Ergashxojayeva Shaxnoza Djasurovna
Toshkent davlat iqtisodiyot universiteti, professor
Saidaxmedova Nodira Ilxomovna
Toshkent davlat iqtisodiyot universiteti, professor
Muxsinov Bekzod Toxirovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shakarov Qulmat Ashirovich
Renessans ta'lim universiteti,dotsent.
Jumaeva Zulfiya Qayumovna
Buxoro Davlat Texnika Universiteti, dotsent.
Norova Salomat Yusupovna
Buxoro Davlat Texnika Universiteti, dotsent.
Avezova Shaxnoza Maxmudjonovna
Buxoro Davlat Texnika Universiteti, dotsent.
Jalilov Jamshid G‘anijonovich
Toshkent davlat iqtisodiyot universiteti, dotsent
Xasanova Gulruh Djumanazarovna
Buxoro Davlat Texnika Universiteti, dotsent.

PEDAGOGIKA FANLARI/ PEDAGOGICAL SCIENCES / ПЕДАГОГИЧЕСКИЕ НАУКИ

Siddiqova Sadoqat G‘afforovna
Buxoro Davlat Texnika Universiteti, dotsent.
Zoyirov Erkin Xalilovich
Buxoro Davlat Texnika Universiteti, DSc, professor
Gaffarov Feruz Hasanovich
Puchon Universiteti, *professor*
Ruziyev Davron Yuldashevich
Buxoro Davlat Universiteti, professor
Muradova Firuza Rashidovna
Buxoro Davlat Texnika Universiteti, professor.
Yunusova Gulandon Samiyevna
Buxoro Davlat Texnika Universiteti, professor.
To‘xtayeva Zebo Sharifovna
Buxoro Davlat Texnika Universiteti, professor.
Navro‘zova Gulchixra Nigmatovna
Buxoro Davlat Texnika Universiteti, dotsent.
Uzoqova Laylo Polvonovna
Buxoro muhandislik – texnologiya instituti, dotsent.
To‘raqulova Marjona Qiyom qizi
Buxoro Davlat Texnika Universiteti, dotsent.
Muradova Zarina Rashidovna
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Sanjar Aslonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shodiyev Jahongir Jo‘raqulovich
Buxoro Davlat Texnika Universiteti, dotsent
Gaffarov Laziz Hasanovich
Buxoro muhandislik – texnologiya instituti, dotsent.

Tahririyat a’zolarning to‘liq ro‘yxati jurnalning www.devos.uz saytida keltirilgan.

Tahririyat hay’ati a’zolarining tarkibini o‘zgartirish hamda qo‘shimcha kiritish yoki tahririyat a’zoligidan chiqarish bosh muharrir tomonidan amalga oshiriladi.

ISSN 3030-3907

Development of science ilmiy jurnali 2024 yil 15 aprel kuni №256445 sonli guvohnoma bilan Davlat ro‘yxatidan o‘tgan.

To‘plam elektron ko‘rinishda (pdf formatida) mualliflarga taqdim etiladi. To‘plamga kiritiladigan maqolalarning mazmuni, ulardagi statistik ma’lumotlar va me’yoriy hujjatlarning to‘g‘riligiga hamda har qanday fikr - mulohazalarga mualliflarning o‘zlari mas’ul hisoblanadilar. Talablarga javob bermaydigan maqolalar to‘plamga kiritilmaydi. Tashkiliy qo‘mita maqola matnini qisqartirish, qisman tuzatish kiritish, seksiyalarga joylashtirish huquqiga ega.

ISSN 3030-3907

UDC: 004.65, 8, 81’2, 81’28

INCORPORATING THE KHOREZM DIALECT INTO UZBEK ASR: AN ANALYTICAL APPROACH TO THE EVOLUTION OF SPEECH RECOGNITION SYSTEMS AND RESOURCES

Kurbonova Ruzikajon Ulug‘bek qizi
PhD student, Urgench State University
E-mail:ruzikajonqurbanova0809@gmail.com

Annotatsiya. Mazkur maqolada nutqni avtomatik tanib olish (ASR) tizimlarining rivojlanish bosqichlari tahlil qilinib, ularni o‘zbek tilining shevalariga, xususan, Xorazm shevasiga joriy etish masalalari yoritiladi. Unda dastlab ASR texnologiyalarining jahon miqyosidagi taraqqiyoti — statistik yondashuvlardan tortib to Wav2Vec 2.0, Conformer, Whisper va Samba-ASR kabi zamonaviy neyron tarmoqlarigacha bo‘lgan evolyutsiyasi ko‘rib chiqiladi. Shuningdek, maqolada adabiy o‘zbek tili uchun yaratilgan ASR resurslari (USC, UzBERT, UzLM, FeruzaSpeech va boshqalar) tahlil qilinib, bu sohadagi yutuqlar e'tirof etiladi. Shu bilan birga, o‘zbek tilining past resursli til sifatidagi xususiyatlari va ochiq nutq korpuslarida fonetik hamda leksik jihatdan standart tildan keskin farq qiluvchi Xorazm shevasi bo‘yicha ma’lumotlarning yetishmasligi muammosi alohida ko‘rsatib o‘tiladi. Tahlil natijalari shuni ko‘rsatadiki, zamonaviy ASR arxitekturalarining o‘zbek tilida to‘laqonli va xatosiz ishlashi uchun Xorazm shevasini mazkur tizimlarga integratsiya qilish o‘ta muhimdir. Maqolada kelgusida aynan Xorazm shevasi uchun maxsus parallel korpuslar yaratish hamda mavjud ilg‘or neyron modellarni ushbu sheva asosida moslashtirib o‘qitish (fine-tuning) sohaning eng dolzarb vazifasi ekanligi asoslab beriladi.

Kalit so‘zlar: Nutqni avtomatik tanib olish (ASR), Xorazm shevasi, o‘zbek tili, parallel nutq korpusi, past resursli tillar, neyron modellar, Whisper, USC.

Аннотация. В данной статье анализируются этапы развития систем автоматического распознавания речи (ASR) и рассматриваются вопросы их внедрения для диалектов узбекского языка, в частности, хорезмского диалекта. В начале рассматривается глобальная эволюция технологий ASR — от статистических подходов до современных архитектур нейронных сетей, таких как Wav2Vec 2.0, Conformer, Whisper и Samba-ASR. Кроме того, в статье анализируются ASR-ресурсы, созданные для литературного узбекского языка (USC, UzBERT, UzLM, FeruzaSpeech и др.), и признаются достижения в этой области. Вместе с тем особо отмечаются статус узбекского языка как малоресурсного (low-resource language) и проблема нехватки данных в открытых речевых корпусах именно по хорезмскому диалекту, который фонетически и лексически резко отличается от стандартного языка. Результаты анализа показывают, что для полноценной

ISSN 3030-3907

и безошибочной работы современных архитектур ASR на узбекском языке интеграция хорезмского диалекта в данные системы имеет критическое значение. В статье обосновывается, что создание специальных параллельных корпусов для хорезмского диалекта, а также дообучение (fine-tuning) существующих передовых нейронных моделей на основе этого диалекта в будущем станут наиболее актуальными задачами в данной сфере.
Ключевые слова: Автоматическое распознавание речи (ASR), хорезмский диалект, узбекский язык, параллельный речевой корпус, малоресурсные языки, нейронные модели, Whisper, USC.
Abstract. This article analyzes the developmental stages of automatic speech recognition (ASR) systems and explores the issues of their integration into Uzbek dialects, specifically focusing on the Khorezm dialect. Initially, the global evolution of ASR technologies is examined, ranging from statistical approaches to modern neural network architectures such as Wav2Vec 2.0, Conformer, Whisper, and Samba-ASR. Furthermore, the article analyzes existing ASR resources developed for the standard literary Uzbek language (USC, UzBERT, UzLM, FeruzaSpeech, etc.) and acknowledges the achievements in this domain. Simultaneously, it highlights the characteristics of the Uzbek language as a low-resource language and explicitly addresses the lack of data in open speech corpora regarding the Khorezm dialect, which significantly differs from the standard language both phonetically and lexically. The analysis results indicate that integrating the Khorezm dialect into these systems is crucial for the comprehensive and accurate performance of modern ASR architectures in the Uzbek language. The article substantiates that the most pressing task in the field is the future development of dedicated parallel corpora for the Khorezm dialect, alongside fine-tuning existing advanced neural models based on this dialectal data.
Keywords: Automatic speech recognition (ASR), Khorezm dialect, Uzbek language, parallel speech corpus, low-resource languages, neural models, Whisper, USC.

INTRODUCTION.

The landscape of speech processing has radically shifted over the past decade. The transition from classical statistical methods—such as Hidden Markov Models (HMMs)—to self-supervised neural architectures like Wav2Vec 2.0, Whisper, and Conformer has solved many traditional bottlenecks in speech recognition. Today, these models perform with remarkable accuracy and adaptability. However, this global success relies heavily on massive amounts of annotated data, a luxury predominantly reserved for high-resource languages.

ISSN 3030-3907

For structurally complex and low-resource languages like Uzbek, the developmental trajectory is fundamentally different. While recent years have seen foundational progress in the Uzbek ASR landscape—marked by the release of speech datasets like USC and FeruzaSpeech, alongside language models such as UzLM and UzBERT—these resources almost exclusively capture the standardized literary language. Consequently, while the basic infrastructure for Uzbek ASR has been laid, the models remain highly rigid.

This rigidity exposes a critical gap in current research: real-world spoken Uzbek is deeply dialectal. Standardized ASR systems frequently fail when processing regional speech patterns, and this breakdown is most pronounced with the Khorezm dialect. Because Khorezmian exhibits profound phonetic, morphological, and lexical divergence from literary Uzbek, it acts almost as a distinct linguistic layer. Acoustic models trained strictly on standard corpora struggle to decode Khorezmian speech, rendering current ASR applications ineffective for a large demographic.

Overcoming this limitation requires moving beyond merely scaling up standard data. It demands a targeted, dialect-aware approach. Thus, rather than presenting a generic overview of ASR history, this paper analyzes the evolution of speech recognition architectures and Uzbek resources with a specific objective: to evaluate how modern neural frameworks can be adapted for the Khorezm dialect. By critically assessing the transition from statistical to deep learning paradigms, we identify the exact methodological gaps in the current Uzbek resource base. Ultimately, this study demonstrates that building dialect-to-standard parallel corpora and fine-tuning state-of-the-art models for the Khorezm dialect is not just the next logical step, but an urgent necessity for the practical evolution of Uzbek ASR.

LITERATURE REVIEW.

Automatic Speech Recognition (ASR) has evolved through three major paradigms. Statistical Hidden Markov Models dominated the field for several decades before being supplanted by deep learning in the 2010s, when the Connectionist Temporal Classification algorithm[1] made it possible to train neural networks directly on unsegmented sequence data and enabled End-to-End architectures. The most recent shift, driven by self-supervised learning, has been particularly transformative for low-resource languages: Wav2Vec 2.0[2] demonstrated that latent representations learned from unlabelled audio can rival fully supervised systems with minimal transcribed data, and its multilingual successor XLSR-53[3] extended this capability across 53 languages. This trajectory culminated in robust large-scale models such as Whisper[4], trained on

ISSN 3030-3907

680,000 hours of weakly supervised multilingual data, which has become the de facto backbone for cross-lingual ASR adaptation.

Uzbek ASR research gained momentum considerably later than for high-resource languages, constrained by its low-resource status and the absence of standardised open datasets. The foundational milestone was the release of the Uzbek Speech Corpus (USC) by Musaev et al.[5], which provided 105 hours of transcribed speech from 958 speakers. This release catalysed subsequent work, most notably the hybrid CTC-attention system of Mukhamadiyev[6], which achieved 14.3% WER on a 207-hour Uzbek dataset. Parallel progress in textual resources improved language modelling, beginning with UzBERT[7] — the first monolingual Uzbek BERT, trained on a 142-million-word news corpus. More recently, Povey[8] released FeruzaSpeech, a 60-hour single-speaker corpus with both Cyrillic and Latin transcripts that complements existing Uzbek datasets for ASR fine-tuning.

Despite this rapid accumulation of resources, the contemporary Uzbek ASR ecosystem is almost exclusively optimised for the standardised literary norm. International experience consistently shows that systems trained only on standard varieties suffer severe degradation when processing regional dialects: the Multi-Genre Broadcast[9] Challenge for Arabic empirically demonstrated how phonetic and lexical divergence across dialects substantially impairs ASR performance even with 1,200 hours of broadcast data. Within the Turkic family, Mussakhojayeva[10] showed that joint multilingual end-to-end training across Kazakh, Bashkir, Kyrgyz, Sakha, and Tatar yields measurable gains for low-resource agglutinative languages but does not, by itself, resolve dialect-internal variation.

The Khorezm dialect of Uzbek represents precisely such a case. Phonologically, it preserves a length distinction in native vowels that has been lost in the standard literary variety[11], and it exhibits prosthetic vowels before initial liquids, distinct vocalic harmony patterns, and dialect-specific suffixes[12]. Lexically and morphologically, it diverges sharply from the Tashkent-based literary norm captured by USC and FeruzaSpeech. Consequently, contemporary Uzbek ASR models are systematically mismatched against Khorezmian input.

In low-resource and highly divergent linguistic environments, training acoustic models from scratch is rarely competitive; transfer learning through targeted fine-tuning of large self-supervised backbones such as Wav2Vec 2.0, XLSR, and Whisper has emerged as the de facto solution[3]. Building on these findings, the present study argues that the most urgent step for closing the standard-versus-dialect gap in Uzbek ASR is the construction of a dedicated parallel corpus of approximately 30,000 Khorezmian sentences paired with their

ISSN 3030-3907

standard literary equivalents, and the use of this corpus to fine-tune state-of-the-art pre-trained acoustic and language models[13].

RESEARCH METHODOLOGY

Addressing the dialectal gap in Uzbek ASR requires a ground-up approach to resource creation and acoustic modeling. Because existing pre-trained systems fail to decode Khorezmian speech accurately, the research methodology is structured into two primary phases: the construction of a highly structured parallel audio-text corpus and the subsequent fine-tuning of advanced neural architectures on this data.

3.1. Corpus Architecture and Data Collection

The foundation of this research is the development of a targeted, 5,000-sentence comparative dataset. To ensure seamless integration with deep learning frameworks, the corpus employs a hybrid storage architecture consisting of a master metadata registry and isolated utterance-level files.

The master registry is maintained in a tabular format (Excel/CSV) to organize the semantic and structural relationships of the data. It is strictly categorized into four primary columns:

- id: A unique numerical identifier for each utterance, serving as the primary key for data alignment.
- domain: The thematic category of the text (e.g., *yangilik* / news). This allows for domain-balanced training and future context-specific evaluation.
- st_Uzbek_Lang: The standard literary Uzbek baseline text, acting as the normative ground truth.
- dialect_Kharezm: The precise phonetic and lexical transcription of the Khorezmian dialect, capturing region-specific morphological shifts.

Simultaneously, the physical data is structured for programmatic access. The standard and dialectal transcriptions from the registry are exported into separate, short .txt files. Each dialectal .txt file is directly mapped via its unique id to a corresponding audio file, recorded by native speakers of the Khorezm dialect. This ensures a strict one-to-one mapping between the spoken acoustic signal and its dialectal transcription.

3.2. Data Processing Pipeline

To prepare the raw paired data for acoustic model training, a rigorous preprocessing pipeline is executed using Python-based scripts.

- Audio Preprocessing: All corresponding audio files are standardized to a 16 kHz sampling rate and single-channel (mono) format, which is the mandatory input specification for modern ASR architectures like Wav2Vec 2.0 and Whisper. Background noise is reduced, and bounding silences are truncated.

ISSN 3030-3907

• Text Normalization: The individual .txt files undergo programmatic cleaning. This includes expanding abbreviations, lowercasing, converting numerical figures into written words, and removing non-speech metadata. The cleaned text files and their paired .wav files are then compiled into a standardized dataset dictionary.

3.3. Acoustic Model Fine-Tuning Setup

Training a deep learning ASR system from scratch is computationally inefficient for a targeted 5,000-sentence dialectal dataset. Therefore, this study employs a Transfer Learning methodology. We leverage robust, pre-trained base models—such as Wav2Vec 2.0 or Whisper—which possess generalized acoustic representations learned from massive multilingual datasets.

During the fine-tuning phase, the model is trained exclusively on the mapped Khorezmian audio and its corresponding dialect_Kharezm .txt files. The model weights are updated to specifically map Khorezmian acoustic features (e.g., distinct vowel harmonies and consonant assimilations) to their accurate textual representations.

3.4. Evaluation Metrics

The performance of the fine-tuned dialectal model will be quantitatively evaluated against baseline standard-Uzbek models using standard industry metrics:

• Word Error Rate (WER): To measure the overall transcription accuracy at the word level.

• Character Error Rate (CER): To capture fine-grained phonetic and morphological recognition errors, which is critical for evaluating the highly agglutinative nature of Khorezmian speech.

ISSN 3030-3907

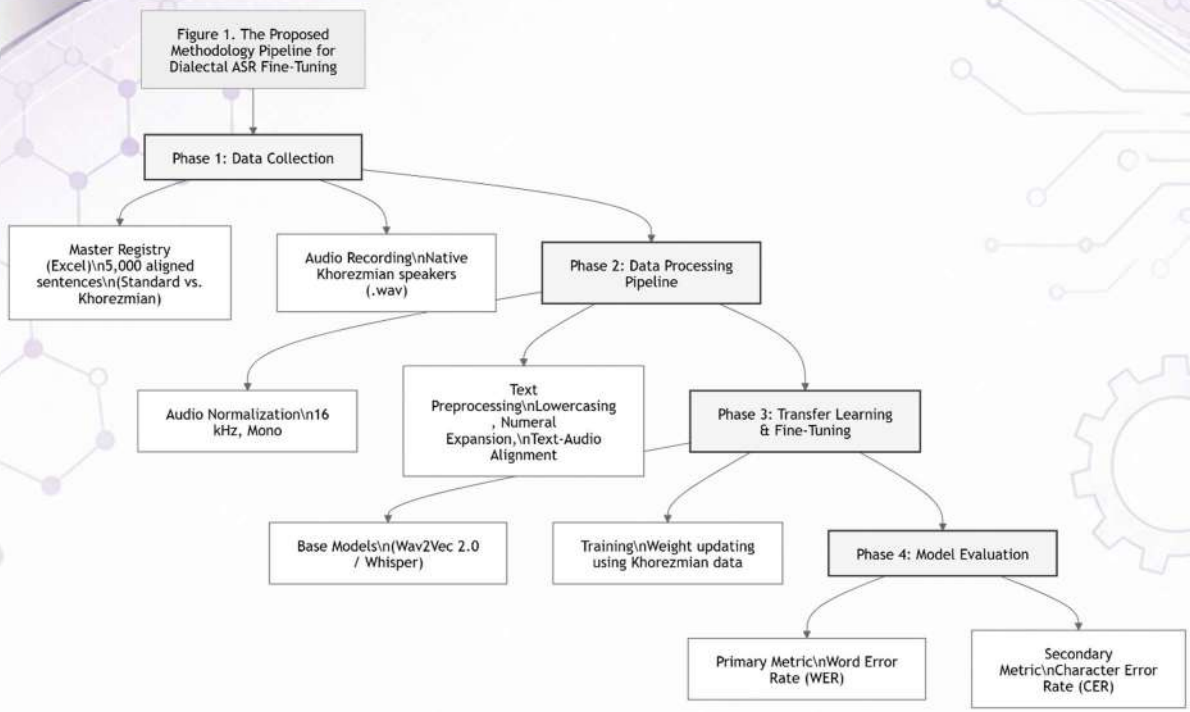


Figure 1. Overview of the proposed methodology pipeline for dialectal automatic speech recognition (ASR) fine-tuning.

Dataset Analysis and Evaluation Results

Since the primary contribution of the current phase of this research is the construction of a foundational dialectal resource, this section analyzes the linguistic properties of the collected dataset and outlines the rigorous evaluation framework that will be applied during the acoustic modeling phase.

4.1. Linguistic and Statistical Analysis of the Corpus The compiled parallel corpus consists of 5,000 paired utterances, explicitly designed to highlight the divergence between standard literary Uzbek and the Khorezmian dialect. A preliminary linguistic analysis of the dataset reveals significant lexical and morphological variance, which justifies the necessity of Transfer Learning.

Key dialectal features captured in the dataset include:

- **Lexical Replacement:** A substantial portion of the corpus contains regional vocabulary that has no acoustic overlap with standard Uzbek (e.g., standard *"guvohnomasiz"* vs. dialectal *"dokumentsiz"*). This out-of-vocabulary (OOV) density is the primary cause of baseline model failure.
- **Morphological Shifts (Oghuz vs. Karluk):** The dialectal text explicitly captures regional verb conjugations and case suffixes (e.g., *"urip ketti"* instead of *"urib yubordi"*).
- **Phonetic Assimilation:** The text preprocessing aligns with acoustic recordings that demonstrate heavy consonant voicing and vowel harmony variations (e.g., *"Tashkentta"* vs. *"Toshkentda"*).

ISSN 3030-3907

By capturing these specific deviations across 5,000 diverse utterances, the dataset provides a sufficiently dense acoustic and textual environment for a neural network to map regional speech patterns.

4.2. Planned Acoustic Evaluation Framework Once the dataset is fully integrated into the fine-tuning pipeline using base models such as Wav2Vec 2.0 or Whisper, the performance of the system will be evaluated using a strict train-test split. To ensure the model learns generalized dialectal rules rather than memorizing the training data, 10% of the corpus will be held out as an unseen test set.

The acoustic performance will be benchmarked against a baseline model (trained solely on standard Uzbek) using two primary metrics:

1. **Word Error Rate (WER):** This will quantify the model's ability to correctly transcribe complete dialectal words and recognize region-specific vocabulary. We hypothesize that the fine-tuned model will demonstrate a sharp reduction in WER compared to the baseline.

2. **Character Error Rate (CER):** Given the highly agglutinative nature of the Uzbek language, CER will serve as a crucial metric for evaluating the model's precision in capturing fine-grained morphological suffixes and phonetic shifts specific to Khorezmian.

4.3. Expected Outcomes and Challenges The anticipated outcome of this methodology is a robust, dialect-aware ASR model capable of seamlessly transcribing Khorezmian speech. However, a key computational challenge remains in balancing the model's new dialectal weights without causing "catastrophic forgetting" of the standard Uzbek language. Future iterations of this research will focus on optimizing the hyper-parameters during Transfer Learning to ensure the model remains highly accurate across both standard and dialectal inputs.

CONCLUSIONS.

This study addressed a critical bottleneck in the evolution of Uzbek Automatic Speech Recognition (ASR): the dialectal gap. While significant progress has been made in developing acoustic and text resources for standard literary Uzbek, the profound linguistic divergence of regional dialects—particularly the Oghuz-based Khorezmian dialect—renders existing baseline models largely ineffective for a substantial demographic.

To systematically bridge this gap, we introduced a foundational methodology centered on targeted resource creation and Transfer Learning. The primary contribution of this research phase is the development of a highly structured, 5,000-utterance parallel text and audio corpus. By meticulously mapping standard Uzbek sentences to their exact Khorezmian equivalents and

ISSN 3030-3907

capturing the genuine acoustic nuances of native speakers, we have established the essential data infrastructure required for dialect-aware neural modeling. This dataset explicitly isolates and addresses the out-of-vocabulary (OOV) challenges, phonetic assimilations, and morphological shifts that traditionally disrupt standardized ASR systems.

Moving forward, the application of this comparative dataset to fine-tune state-of-the-art pre-trained architectures (such as Wav2Vec 2.0 or Whisper) represents the immediate next step. We anticipate that this dialect-specific fine-tuning will drastically reduce Word Error Rates (WER) for regional speech. Future research will focus on optimizing hyper-parameters to balance the model's performance, ensuring that the network retains its proficiency in standard Uzbek while accurately decoding Khorezmian inputs—effectively preventing catastrophic forgetting.

Ultimately, integrating regional dialects into modern neural architectures is not merely a theoretical exercise, but an urgent necessity. Overcoming the dialectal gap ensures equitable human-computer interaction, enhances the robustness of voice-activated technologies, and expands digital accessibility for all linguistic demographics within Uzbekistan.

REFERENCES

1. A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, "Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks," in *Proceedings of the 23rd International Conference on Machine Learning (ICML)*, ACM, 2006, pp. 369–376. doi: 10.1145/1143844.1143891.

2. A. Baevski, H. Zhou, A. Mohamed, and M. Auli, "wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2020, pp. 12449–12460.

3. A. Conneau, A. Baevski, R. Collobert, A. Mohamed, and M. Auli, "Unsupervised Cross-Lingual Representation Learning for Speech Recognition," in *Proceedings of Interspeech 2021*, 2021, pp. 2426–2430. doi: 10.21437/Interspeech.2021-329.

4. A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust Speech Recognition via Large-Scale Weak Supervision," in *Proceedings of the 40th International Conference on Machine Learning (ICML)*, 2023, pp. 28492–28518.

5. M. Musaev, I. Khujayorov, and M. Ochilov, "USC: An Open-Source Uzbek Speech Corpus and Initial Speech Recognition Experiments," in *Speech and Computer (SPECOM 2021)*, in Lecture Notes in Computer Science, vol. 12997. Springer, 2021, pp. 437–447. doi: 10.1007/978-3-030-87802-3_40.

ISSN 3030-3907

6. A. Mukhamadiyev, I. Khujayarov, O. Djuraev, and J. Cho, "Automatic Speech Recognition Method Based on Deep Learning Approaches for the Uzbek Language," *Sensors*, vol. 22, no. 10, p. 3683, 2022, doi: 10.3390/s22103683.
7. B. Mansurov and A. Mansurov, "UzBERT: pretraining a BERT model for Uzbek," 2021. [Online]. Available: https://huggingface.co/coppercitylabs/
8. A. Povey and K. Povey, "FeruzaSpeech: A 60 Hour Uzbek Read Speech Corpus with Punctuation, Casing, and Context," Apr. 2024. doi: 10.48550/arXiv.2410.00035.
9. A. Ali *et al.*, "The MGB-2 Challenge: Arabic Multi-Dialect Broadcast Media Recognition," in *2016 IEEE Spoken Language Technology Workshop (SLT)*, IEEE, 2016, pp. 279–284. doi: 10.1109/SLT.2016.7846277.
10. S. Mussakhojayeva, Y. Khassanov, and H. A. Varol, "Multilingual End-to-End ASR for Low-Resource Turkic Languages with Common Alphabets," *Sci. Rep.*, vol. 14, p. 13632, 2024, doi: 10.1038/s41598-024-64848-1.
11. A. G. McCollum and S. Chen, "Uzbek," *J. Int. Phon. Assoc.*, vol. 51, no. 3, pp. 1–25, 2021, doi: 10.1017/S0025100319000264.
12. Z. Dosimov and S. Alimova, "The Vocabulary Layer of the Khorezm Kipchak Dialect," *Psychology and Education Journal*, vol. 58, no. 1, 2021.
13. M. Sharipov, J. Mattiev, J. Sobirov, and R. Baltayev, "Creating a morphological and syntactic tagged corpus for the Uzbek language," *arXiv preprint arXiv:2210.15234*, 2022.

ISSN 3030-3907