

ANNOTATION
Abduali Balzhan Abdirasilkyzy
for the dissertation prepared for the degree
of Doctor of Philosophy (PhD)
on the educational program
8D07109 – “Automation and Internet of Things”
on the topic “ Development and research of a model and tools for an
automated system for generating meeting transcripts for the Kyrgyz–Kazakh
language pair”

Relevance of the Research Topic. Currently, the rapid development of digital technologies is increasing the importance of natural language processing, automatic speech recognition, and machine translation. Nevertheless, Turkic languages such as Kyrgyz and Kazakh belong to the category of low-resource languages. The lack of annotated audio data, parallel texts, and high-quality corpora for these languages significantly limits the effective development of modern intelligent systems.

Although the Kyrgyz–Kazakh language pair is linguistically close, speech processing, automatic translation, and text summarization systems for this direction remain insufficiently studied. The agglutinative nature of Turkic languages, their morphological complexity, and the large variety of word forms complicate the process of building robust language models.

At the same time, the demand for automatic generation of meeting and conference transcripts is continuously increasing in modern society. Such systems involve several complex stages, including speech-to-text conversion, text processing, extraction of key information, and summarization. In multilingual environments, these solutions can significantly accelerate information exchange and reduce human effort.

In this regard, the development and investigation of models and tools for automatic meeting transcript generation for the Kyrgyz–Kazakh language pair represents an important scientific and practical task.

Today, artificial intelligence technologies are widely used in speech processing, text analysis, and automatic documentation systems. In particular, automatic speech recognition, machine translation, and meeting transcript automation have become highly important research areas. These technologies play a significant role in education, public administration, business, media, international communication, and the development of digital services. The demand for multilingual systems is growing rapidly, making the study of effective models for relatively low-resource languages such as Kazakh increasingly relevant.

Automatic speech recognition systems enable the conversion of spoken language into text. Such systems are applied in audio archive digitization, subtitle generation, assistive technologies, call-center systems, and automatic meeting recording. In recent years, models such as Whisper, MMS, and Turkic ASR have attracted considerable attention. Their main advantages include multilingual capabilities, adaptation to low-resource languages, and improved speech recognition quality in real-world conditions.

The field of machine translation is also developing rapidly. Earlier translation systems were mainly limited to specific language pairs, whereas modern neural models can perform translation across dozens or even hundreds of languages. In this context, modern models such as mT5, Gemma, and NLLB are of particular interest. These models are capable not only of translation but also of multilingual text understanding, processing, and generation. Therefore, comparing their architectures, operational principles, and application domains is scientifically and practically significant.

In addition, specialized evaluation metrics are used to assess the quality of speech recognition and machine translation systems. Among them, BLEU, chrF, WER, and TER are the most widely used. These metrics compare system outputs with reference texts and provide quantitative evaluation of error rates, similarity levels, and translation or recognition quality. Proper understanding of these metrics ensures the objectivity of the research and helps clearly identify the strengths and limitations of the models.

Furthermore, automatic meeting transcript generation has recently become an important research direction. Such systems not only convert speech into text but also automatically identify key ideas, decisions, tasks, and conclusions discussed during meetings. This provides significant advantages in reducing time consumption, simplifying document management, and organizing information within institutions and organizations. However, research in this area for the Kazakh language remains limited, making it important to collect, analyze, and compare existing studies.

The purpose of this work is to investigate Whisper, Turkic ASR, and MMS models for speech recognition, as well as mT5, Gemma, and NLLB models for machine translation; to describe their mathematical foundations; to review related scientific studies; and to explain BLEU, chrF, WER, and TER evaluation metrics. In addition, the work aims to analyze modern studies on meeting transcript automation and demonstrate the development prospects of this field for the Kazakh language.

Purpose of the Dissertation Work

The purpose of this dissertation is to develop and investigate an intelligent system model for the automatic generation of meeting transcripts for the Kyrgyz–Kazakh language pair.

Research Objectives

To achieve the research goal, the following objectives are defined:

- To investigate an automatic speech recognition (ASR) model for converting Kyrgyz meeting speech into text;
- To develop a CSE-based Morphology-Aware ASR Post-Processing method;
- To investigate a neural machine translation model for translating text from Kyrgyz into Kazakh;
- To develop an integrated ASR → MT system for automated meeting transcript generation;
- To conduct experimental evaluation of the developed models.

Research Object. The process of transforming Kyrgyz speech into Kazakh text.

Research Subject. Methods and models of speech recognition and machine translation in a multilingual environment.

Scientific Novelty of the Research

- A large parallel corpus for the Kyrgyz-Kazakh language pair was developed, containing 782,293 parallel sentences;
- The Morphology-Aware Post-Processing method based on CSE technology was proposed to correct the results of Kyrgyz speech recognition;
- A system for automatically generating meeting transcripts for the Kyrgyz-Kazakh language pair has been developed.

Statements Submitted for Defense

- development of a large parallel corpus of 782,293 parallel sentences for the Kyrgyz-Kazakh language pair;
- Morphology method Aware Post Processing based on CSE technology for error correction in Kyrgyz speech recognition results;
- development of a system for automatic generation of meeting transcripts for the Kyrgyz-Kazakh language pair.

Theoretical and Practical Significance of the Work

- Enables automation of meeting transcript generation in government institutions, educational organizations, and business environments;
- Contributes to the development of speech technologies for the Kyrgyz and Kazakh languages;
- Provides a universal solution applicable to multilingual intelligent systems;
- Serves as a foundation for developing effective software tools for low-resource languages.

Relationship with Scientific Research Programs

The dissertation work was carried out within the framework of the grant-funded research project AP23487816 “Study of neural models for the formation of transcripts of speech and minutes of meetings in Turkic languages” (2024–2026).

Scope and Structure of the Dissertation

The dissertation consists of an introduction, three chapters, a conclusion, and a list of references.

The first chapter discusses the theoretical foundations of automatic speech recognition (ASR) and machine translation systems for the Kyrgyz–Kazakh language pair. An analysis of modern scientific studies on ASR technologies for the Kyrgyz and Kazakh languages, neural machine translation approaches, and automated transcript generation systems was conducted.

The second chapter presents the methodology for developing an

automated transcript generation system for the Kyrgyz–Kazakh language pair. The client–server architecture of the system, methods for data collection and processing, stages of speech recognition model investigation, the Morphology-Aware Post-Processing method, machine translation model development stages, and models for transcript generation are described.

The third chapter presents the experimental results and analysis of the proposed automated system. The quality of ASR models was evaluated, and the effectiveness of the Morphology-Aware Post-Processing method was analyzed. In addition, the results of Kyrgyz-to-Kazakh translation were presented, and the proposed system was compared with existing approaches. Practical application features, system limitations, and future development directions were also discussed. The effectiveness of the proposed approach for the low-resource Kyrgyz–Kazakh language pair was substantiated, and further development opportunities for the automated transcript generation system were described.

Publications

Six scientific publications were published based on the dissertation research, including three papers in peer-reviewed journals indexed in international citation databases (“Web of Science” and/or “Scopus”) and one paper published in international conference proceedings.

In addition, based on the results of the research work, a copyright registration certificate for the “Kazakh–Kyrgyz Parallel Text Corpus ‘KazKyrParC’” was officially registered in the State Register of Rights to Objects Protected by Copyright of the Republic of Kazakhstan under certificate No. 62995 dated October 14, 2025 (Appendix).

Publications:

- U. Tukeyev, A. Shormakova, A. Karibayeva, D. Rakhimova, B. Abduali, D. Amirova, N. Rakhmanberdi, R. Aliyev. An Integrated Approach to Adapting Open-Source AI Models for Machine Translation of Low-Resource Turkic Languages 2026 Computers 2026, MDPI, 15(2), 73. (Percentile 84, Q1) <https://doi.org/10.3390/computers15020073>
- A. Karibayeva, V. Karyukin, B. Abduali, D. Amirova, Speech Recognition and Synthesis Models and Platforms for the Kazakh Language 2025 Information 2025, MDPI, 16(10), 879; (Percentile 74, Q2) <https://doi.org/10.3390/info16100879>
- A. Karibayeva, V. Karyukin, U. Tukeyev, B. Abduali, D. Amirova, D. Rakhimova, R. Aliyev, A. Shormakova. The Development and Experimental Evaluation of a Multilingual Speech Corpus for Low-Resource Turkic Languages 2025 Applied Science. MDPI, 2025, 15(24), 12880; (Percentile 79, Q2) <https://doi.org/10.3390/app152412880>
- Abduali B., Milosz, M., Tukeyev, U., Karibayeva, A. Methodology for creating datasets of parallel sentences in low-resource languages by using AI. International Journal of Innovative Research and Scientific Studies, 8(9), 13–23. 2025 <https://doi.org/10.53894/ijirss.v8i9.10605>

- Abduali B., Tukeyev U., Zhumanov Zh., Israilova N. Study of Kyrgyz-Kazakh Neural Machine Translation 2025 17th Asian Conference on Intelligent Information and Database Systems, proceedings conference paper (Scopus) Doi: 10.1007/978-981-96-5881-7_21.
- A. Karibaeva, B. Abduali, D. Amirova. Analysis of reverse translation technology in English-Kazakh neural machine translation, 31.10-1.11.2024, IX International Scientific and Practical Conference "Informatics and Applied Mathematics"
- Copyright certificate «Kazakh–Kyrgyz parallel corpus of texts «KazKyrParC»» № №62995. 14.10.2025