ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

№2
2026

ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

CENTRAL ASIAN ACADEMIC
RESEARCH CENTER

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

2 (358)

APRIL – JUNE 2026

PUBLISHED SINCE JANUARY 1963
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

Chief Editor:

MUTANOV Galimkair Mutanovich, doctor of technical sciences, professor, academician of NAS RK, (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=6506682964, https://www.webofscience.com/wos/author/record/1423665

EDITORIAL BOARD:

KALIMOLDAYEV Maksat Nuradilovich, (Deputy Editor-in-Chief), Doctor of Physical and Mathematical Sciences, Professor, Academician of NAS RK, Advisor to the General Director of the Institute of Information and Computing Technologies of the CS MES RK, Head of the Laboratory (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=56153126500, https://www.webofscience.com/wos/author/record/2428551

MAMYRBAEV Orken Zhumazhanovich, (Academic Secretary), PhD in Information Systems, Deputy Director for Science of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=55967630400, https://www.webofscience.com/wos/author/record/1774027

BAIGUNCHEKOV Zhumadil Zhanabaevich, Doctor of Technical Sciences, Professor, Academician of NAS RK, Institute of Cybernetics and Information Technologies, Department of Applied Mechanics and Engineering Graphics, Satbayev University (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=6506823633, https://www.webofscience.com/wos/author/record/1923423

WOICIK Waldemar, Doctor of Technical Sciences (Phys.-Math.), Professor of the Lublin University of Technology (Lublin, Poland), https://www.scopus.com/authid/detail.uri?authorId=7005121594, https://www.webofscience.com/wos/author/record/678586

SMOLARJ Andrej, Associate Professor Faculty of Electronics, Lublin polytechnic university (Lublin, Poland), https://www.scopus.com/authid/detail.uri?authorId=56249263000, https://www.webofscience.com/wos/author/record/1268523

KEILAN Alimkhan, Doctor of Technical Sciences, Professor (Doctor of science (Japan)), chief researcher of Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=8701101900, https://www.webofscience.com/wos/author/record/1436451

KHAIROVA Nina, Doctor of Technical Sciences, Professor, Chief Researcher of the Institute of Information and Computational Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=37461441200, https://www.webofscience.com/wos/author/record/1768515

OTMAN Mohamed, PhD, Professor of Computer Science Department of Communication Technology and Networks, Putra University Malaysia (Selangor, Malaysia), https://www.scopus.com/authid/detail.uri?authorId=56036884700, https://www.webofscience.com/wos/author/record/747649

NYSANBAYEVA Saule Yerkebulanovna, Doctor of Technical Sciences, Associate Professor, Senior Researcher of the Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=55453992600, https://www.webofscience.com/wos/author/record/3802041

USATOVA Olga Alexandrovna, PhD, Associate Professor, Chief Scientific Secretary of the Institute of Information and Computing Technologies of the National Academy of Sciences of the Republic of Kazakhstan (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=57204581062, https://www.webofscience.com/wos/author/record/JCO-3058-2023

KAPALOVA Nursulu Aldazharovna, Candidate of Technical Sciences, Head of the Laboratory cybersecurity, Institute of Information and Computing Technologies CS MES RK (Almaty, Kazakhstan), https://www.scopus.com/authid/detail.uri?authorId=57191242124,

KOVALYOV Alexander Mikhailovich, Doctor of Physical and Mathematical Sciences, Academician of the National Academy of Sciences of Ukraine, Institute of Applied Mathematics and Mechanics (Donetsk, Ukraine), https://www.scopus.com/authid/detail.uri?authorId=7202799321, https://www.webofscience.com/wos/author/record/38481396

MIKHALEVICH Alexander Alexandrovich, Doctor of Technical Sciences, Professor, Academician of the National Academy of Sciences of Belarus (Minsk, Belarus), https://www.scopus.com/authid/detail.uri?authorId=7004159952, https://www.webofscience.com/wos/author/record/46249977

TIGHINEANU Ion Mihailovich, Doctor of Physical and Mathematical Sciences, Academician, President of the Academy of Sciences of Moldova, Technical University of Moldova (Chisinau, Moldova), https://www.scopus.com/authid/detail.uri?authorId=7006315935, https://www.webofscience.com/wos/author/record/524462

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Owner: «Central Asian Academic Research Center» LLP (Almaty).
Certificate **№ KZ77VPY00121154** on the re-registration of the periodical printed and online publication of the information agency, issued on **05.06.2025** by the Republican State Institution «Information Committee» of the Ministry of Culture and Information of the Republic of Kazakhstan
Subject area*: information and communication technologies.*
Currently: *included in the list of journals recommended by the CCSES MSHE RK in the direction of «Information and communication technologies».*
Periodicity: *4 times a year.*
http://www.physico-mathematical.kz/index.php/en/

© «Central Asian Academic Research Center» LLP, 2026

БАС РЕДАКТОР:

МҰТАНОВ Ғалымқайыр Мұтанұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=6506682964, https://www.webofscience.com/wos/author/record/1423665

РЕДАКЦИЯ АЛҚАСЫ:

ҚАЛИМОЛДАЕВ Мақсат Нұрәділұлы, (бас редактордың орынбасары), физика-математика ғылымдарының докторы, профессор, ҚР ҰҒА академигі, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының кеңесшісі, зертхана меңгерушісі (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=56153126500, https://www.webofscience.com/wos/author/record/2428551

МАМЫРБАЕВ Өркен Жұмажанұлы (ғалым хатшы), Ақпараттық жүйелер саласындағы техника ғылымдарының (PhD) докторы, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» директорының ғылым жөніндегі орынбасары (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=55967630400, https://www.webofscience.com/wos/author/record/1774027

БАЙГУНЧЕКОВ Жұмаділ Жаңабайұлы, техника ғылымдарының докторы, профессор, ҚР ҰҒА академигі, Кибернетика және ақпараттық технологиялар институты, Қолданбалы механика және инженерлік графика кафедрасы, Сәтбаев университеті (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=6506823633, https://www.webofscience.com/wos/author/record/1923423

ВОЙЧИК Вальдемар, техника ғылымдарының докторы (физ-мат), Люблин технологиялық университетінің профессоры (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=7005121594, https://www.webofscience.com/wos/author/record/678586

СМОЛАРЖ Анджей, Люблин политехникалық университетінің электроника факультетінің доценті (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=56249263000, https://www.webofscience.com/wos/author/record/1268523

КЕЙЛАН Әлімхан, техника ғылымдарының докторы, профессор (ғылым докторы (Жапония)), ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» бас ғылыми қызметкері (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=8701101900, https://www.webofscience.com/wos/author/record/1436451

ХАЙРОВА Нина, техника ғылымдарының докторы, профессор, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» бас ғылыми қызметкері (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=37461441200, https://www.webofscience.com/wos/author/record/1768515

ОТМАН Мохаммед, PhD, Информатика, Коммуникациялық технологиялар және желілер кафедрасының профессоры, Путра университеті Малайзия (Селангор, Малайзия), https://www.scopus.com/authid/detail.uri?authorId=56036884700, https://www.webofscience.com/wos/author/record/747649

НЫСАНБАЕВА Сауле Еркебұланқызы, техника ғылымдарының докторы, доцент, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институтының» аға ғылыми қызметкері (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=55453992600, https://www.webofscience.com/wos/author/record/3802041

УСАТОВА Ольга Александровна, PhD, қауымдастырылған профессор, ҚР ҒЖБМ "Ақпараттық және есептеу технологиялары институтының" бас ғалым хатшысы (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=57204581062, https://www.webofscience.com/wos/author/record/JCO-3058-2023

КАПАЛОВА Нұрсұлу Алдажарқызы, техника ғылымдарының кандидаты, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты», Киберқауіпсіздік зертханасының меңгерушісі (Алматы, Қазақстан), https://www.scopus.com/authid/detail.uri?authorId=57191242124,

КОВАЛЕВ Александр Михайлович, физика-математика ғылымдарының докторы, Украина Ұлттық Ғылым академиясының академигі, Қолданбалы математика және механика институты (Донецк, Украина), https://www.scopus.com/authid/detail.uri?authorId=7202799321, https://www.webofscience.com/wos/author/record/38481396

МИХАЛЕВИЧ Александр Александрович, техника ғылымдарының докторы, профессор, Беларусь Ұлттық Ғылым академиясының академигі (Минск, Беларусь), https://www.scopus.com/authid/detail.uri?authorId=7004159952, https://www.webofscience.com/wos/author/record/46249977

ТИГИНЯНУ Ион Михайлович, физика-математика ғылымдарының докторы, академик, Молдова Ғылым Академиясының президенті, Молдова техникалық университеті (Кишинев, Молдова), https://www.scopus.com/authid/detail.uri?authorId=7006315935, https://www.webofscience.com/wos/author/record/524462

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Меншіктеуші: «Орталық Азия академиялық ғылыми орталығы» ЖШС (Алматы).
Ақпарат агенттігінің мерзімді баспасөз басылымын, ақпарат агенттігін және желілік басылымды қайта есепке қою туралы ҚР Мәдениет және Ақпарат министрлігі «Ақпарат комитеті» Республикалық мемлекеттік мекемесі **05.06.2025** ж. берген **№ KZ77VPY00121154** Куәлік.
Тақырыптық бағыты: *ақпараттық-коммуникациялық технологиялар*
Қазіргі уақытта: *«ақпараттық-коммуникациялық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*
Мерзімділігі: *жылына 4 рет.*
http://www.physico-mathematical.kz/index.php/en/

© «Орталық Азия академиялық ғылыми орталығы» ЖШС, 2026

Главный редактор:

МУТАНОВ Галимкаир Мутанович, доктор технических наук, профессор, академик НАН РК, (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=6506682964, https://www.webofscience.com/wos/author/record/1423665

Редакционная коллегия:

КАЛИМОЛДАЕВ Максат Нурадилович, (заместитель главного редактора), доктор физико-математических наук, профессор, академик НАН РК, советник генерального директора «Института информационных и вычислительных технологий» КН МНВО РК, заведующий лабораторией (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=56153126500, https://www.webofscience.com/wos/author/record/2428551

МАМЫРБАЕВ Оркен Жумажанович, (ученый секретарь), доктор философии (PhD) по специальности «Информационные системы», заместитель директора по науке РГП «Институт информационных и вычислительных технологий» Комитета науки МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=55967630400, https://www.webofscience.com/wos/author/record/1774027

БАЙГУНЧЕКОВ Жумадил Жанабаевич, доктор технических наук, профессор, академик НАН РК, Институт кибернетики и информационных технологий, кафедра прикладной механики и инженерной графики, Университет Сатпаева (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=6506823633, https://www.webofscience.com/wos/author/record/1923423

ВОЙЧИК Вальдемар, доктор технических наук (физ.-мат.), профессор Люблинского технологического университета (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=7005121594, https://www.webofscience.com/wos/author/record/678586

СМОЛАРЖ Анджей, доцент факультета электроники Люблинского политехнического университета (Люблин, Польша), https://www.scopus.com/authid/detail.uri?authorId=56249263000, https://www.webofscience.com/wos/author/record/1268523

КЕЙЛАН Алимхан, доктор технических наук, профессор (Doctor of science (Japan)), главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=8701101900, https://www.webofscience.com/wos/author/record/1436451

ХАЙРОВА Нина, доктор технических наук, профессор, главный научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=37461441200, https://www.webofscience.com/wos/author/record/1768515

ОТМАН Мохамед, доктор философии, профессор компьютерных наук, Департамент коммуникационных технологий и сетей, Университет Путра Малайзия (Селангор, Малайзия), https://www.scopus.com/authid/detail.uri?authorId=56036884700, https://www.webofscience.com/wos/author/record/747649

НЫСАНБАЕВА Сауле Еркебулановна, доктор технических наук, доцент, старший научный сотрудник РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=55453992600, https://www.webofscience.com/wos/author/record/3802041

УСАТОВА Ольга Александровна, PhD, ассоциированный профессор, Главный ученый секретарь «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=57204581062, https://www.webofscience.com/wos/author/record/JCO-3058-2023

КАПАЛОВА Нурсулу Алдажаровна, кандидат технических наук, заведующий лабораторией кибербезопасности РГП «Института информационных и вычислительных технологий» КН МНВО РК (Алматы, Казахстан), https://www.scopus.com/authid/detail.uri?authorId=57191242124,

КОВАЛЕВ Александр Михайлович, доктор физико-математических наук, академик НАН Украины, Институт прикладной математики и механики (Донецк, Украина), https://www.scopus.com/authid/detail.uri?authorId=7202799321, https://www.webofscience.com/wos/author/record/38481396

МИХАЛЕВИЧ Александр Александрович, доктор технических наук, профессор, академик НАН Беларуси (Минск, Беларусь), https://www.scopus.com/authid/detail.uri?authorId=7004159952, https://www.webofscience.com/wos/author/record/46249977

ТИГИНЯНУ Ион Михайлович, доктор физико-математических наук, академик, президент Академии наук Молдовы, Технический университет Молдовы (Кишинев, Молдова), https://www.scopus.com/authid/detail.uri?authorId=7006315935, https://www.webofscience.com/wos/author/record/524462

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Собственник: *ТОО «Центрально-азиатский академический научный центр» (г. Алматы).*
Свидетельство о постановке на переучет периодического печатного издания, информационного агентства и сетевого издания **№ KZ77VPY00121154.** Дата выдачи **05.06.2025**
Тематическая направленность: *информационно-коммуникационные технологии.*
В настоящее время: *вошел в список журналов, рекомендованных КОКСНВО МНВО РК по направлению «информационно-коммуникационные технологии».*
Периодичность: *4 раза в год.*
http://www.physico-mathematical.kz/index.php/en/

© ТОО «Центрально-азиатский академический научный центр», 2026

CONTENTS

COMPUTER SCIENCE

МАЗМҰНЫ

КОМПЬЮТЕРЛІК ҒЫЛЫМДАР

СОДЕРЖАНИЕ

КОМПЬЮТЕРНЫЕ НАУКИ

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Namber 358 (2026). 414–432

https://doi.org/10.32014/2026.2518-1726.446

IRSTI 28.23.15
UDC 004.89

© Tursynova N.[1,2], Yerimbetova A[1,3], Amangeldy N.[1,2], Zhumabayeva A.[2], Daiyrbayeva E.[1,4*], 2026.
[1]Institute of Information and Computational Technologies, Almaty, Kazakhstan;
[2]L.N. Gumilyov Eurasian National University, Astana, Kazakhstan;
[3]META University, Astana, Kazakhstan;
[4]Satbayev University, Almaty, Kazakhstan.
E-mail: nurbekkyzyelmira@gmail.com

COMPARATIVE ANALYSIS OF MULTILINGUAL TRANSFORMER MODELS FOR KAZAKH-TO-GLOSS TRANSLATION

Tursynova Nazira — Software Engineer, Institute of Information and computing technologies, Almaty, Kazakhstan; PhD student, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: ntursynova000@gmail.com, https://orcid.org/0000-0002-1857-9028;
Yerimbetova Aigerim — Leading researcher, Institute of Information and computing technologies, Almaty, Kazakhstan; PhD, Cand. Tech. Sci., Associate Professor, Research Professor, META University, Astana, Kazakhstan,
E-mail: aigerian8888@gmail.com, https://orcid.org/0000-0002-2013-1513;
Amangeldy Nurzada — Senior Researcher, Institute of Information and computing technologies, Almaty, Kazakhstan; PhD, Researcher, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: amangeldi_n_3@enu.kz , https://orcid.org/0000-0002-4669-9254;
Zhumabayeva Assel — PhD, Senior Lecturer, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: zhumabayeva_as@enu.kz , https://orcid.org/0000-0002-8984-1492;
Daiyrbayeva Elmira — Researcher, Institute of Information and computing technologies, Almaty, Kazakhstan; Master, senior lecturer, Satbayev University, Almaty, Kazakhstan,
E-mail: nurbekkyzyelmira@gmail.com, https://orcid.org/0000-0002-4255-5456.

Abstract. Kazakh sign language is the focus of attention among scientists. In recent decades, increasing attention has been paid to the creation of technologies capable of removing barriers to communication and interaction between people with disabilities and society. The lack of a parallel corpus in sign language opens up the topic of text translation into gloss for urgent and in-depth research. Unfortunately, digital technologies are not accessible to people with hearing and speech impairments. Text-to-gloss translation is still relevant for Kazakh sign language. In particular, because one sign has several meanings, synthesizing it into gloss in a software environment takes a long time. This article discusses the translation of text into gloss using five different models based on a multilingual transformer. That is, it examines whether the models are faster or slower to read and whether they consume more or less resources. The mBART-large-50, M2M100-418M, mT5-base, MarianMT, and T5-base were selected for the experiment. All of these models were trained using the same parameters. The results were evaluated using widely used standard metrics (BLEU, ROUGE-L, and chrF++), and a qualitative analysis of the models was also performed. The study found that larger multilingual models provide more accurate translations. Although the mBART-large-50 and M2M100-418M models performed well, the other smaller multilingual models were faster and less resource-intensive. Translating text into a glossary is still relevant for Kazakh sign language. In particular, because one sign language has several meanings, synthesizing it into a glossary in a programming environment takes a significant amount of time.

Keywords: Kazakh sign language, text translation to gloss, multilingual Transformer platform, BLEU score, training

Financing. This work was carried out with the financial support of the Science Committee of the Ministry of Science and higher education of the Republic of Kazakhstan (Grant No. BR24992875).

For citations: Tursynova N., Yerimbetova A., Amangeldy N., Zhumabayeva A., Daiyrbayeva E. Comparative analysis of multilingual transformer models for Kazakh-to-gloss translation. Academic Scientific Journal of Computer Science, 2026. — No.2. – P. 414-432. DOI https://doi.org/10.32014/2026.2518-1726.446

© Турсынова Н.[1,2], Еримбетова А.[1,3], Амангелді Н.[1,2], Жумабаева А.[2], Дайырбаева Э.[1,4*], 2026.
[1]ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы, Қазақстан;
[2]Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті, Астана, Қазақстан;
[3]META University, Алматы, Қазақстан;
[4] Қ.И. Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан.
E-mail: nurbekkyzyelmira@gmail.com

ҚАЗАҚ ТІЛІНЕН ГЛОССҚА АУДАРУ ҮШІН КӨПТІЛДІ ТРАНСФОРМЕРЛІК МОДЕЛЬДЕРДІҢ САЛЫСТЫРМАЛЫ ТАЛДАУЫ

Турсынова Назира — инженер-бағдарламашы, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», PhD докторант, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,

E-mail: ntursynova000@gmail.com, https://orcid.org/0000-0002-1857-9028;
Еримбетова Айгерим — жетекші ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», PhD, т.ғ.к., қауымдастырылған профессор, профессор – зерттеуші, META University , Алматы, Қазақстан,
E-mail: aigerian8888@gmail.com, https://orcid.org/0000-0002-2013-1513;
Амангелді Нұрзада — аға ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», PhD, зерттеуші, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: amangeldi_n_3@enu.kz , https://orcid.org/0000-0002-4669-9254;
Жумабаева Асель Сагнаевна — PhD, аға оқытушысы, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: zhumabayeva_as@enu.kz , https://orcid.org/0000-0002-8984-1492;
Дайырбаева Эльмира Нурбеккызы – ғылыми қызметкер, ҚР ҒЖБМ ҒК қарасты «Ақпараттық және есептеу технологиялары институты», Алматы, Қазақстан; магистр, аға оқытушы, Қ.И.Сәтпаев атындағы Қазақ ұлттық техникалық зерттеу университеті, Алматы, Қазақстан,
E-mail: nurbekkyzyelmira@gmail.com, https://orcid.org/0000-0002-4255-5456.

Аннотация. Қазақ ым тілі әлі де ғалымдардың назарында. Соңғы онжылдықтарда мүмкіндігі шектеулі жандар мен қоғамның қарым-қатынасы мен өзара әрекеттесуіндегі кедергілерді жоюға қабілетті технологияларды құруға көбірек көңіл бөлінді. Ым тілінде параллельді корпустың аз болуы, мәтінді глоссқа аудару тақырыбы өзекті және терең зерттеуге жол ашады. Себебі есту және сөйлеу қаблетінде шектеу бар адамдар үшін цифрлықық технология өкінішке орай қолжетімсіз болып отыр. Қазақ ым тілі үшін әліде мәтіннен глоссқа аудару өзекті болып тұр. Нақтырақ айтқанда бір ым тілінің бірнеше мағнасы болуына байланысты оны программалау ортасында глоссқа синтездеу айтарлықтай уақытты талап етеді. Бұл мақалада көптілді Transformer-ға негізделген бес түрлі модельдерді қолдану арқылы мәтінің глоссқа аударылуы қарастырылады. Яғни модельдердің жылдамырақ немесе әлсіз оқылатыны және ресуртың көп немесе аз жұмсалуы зерттеледі. Тәжірбиені іске асыру үшін mBART-large-50, M2M100-418M, mT5-base, MarianMT және T5-base модельдері таңдалды. Бұл модельдердің барлығына бірдей параметрлер қолдану арқылы оқытылды. Нәтижелерін кең таралған стандарттар BLEU, ROUGE-L, және chrF++) метрикаларымен бағаланып, модельдерге сапалы талдау жасалды. Зерттеу барысында үлкен көптілді модельдердің дәлірек аударма жасайтыны анықталды. mBART-large-50 және M2M100-418M модельдері жақсы өнімділікті көрсетсе, қалған кішірек көптілді модельдер жылдамырақ оқылады, аз ресурс жұмсайтыны дәлелденді. Қазақ ым тілі үшін әліде мәтіннен глоссқа аудару өзекті болып тұр. Нақтырақ айтқанда бір ым тілінің бірнеше мағынасы болуына байланысты оны программалау ортасында глоссқа синтездеу айтарлықтай уақытты талап етеді.

Түйін сөздер: Қазақ ым тілі, мәтінді глоссқа аудару, көптілді Transformer негізі, BLEU бағалау, оқыту

© Турсынова Н.[1,2], Еримбетова А.[1,3], Амангелді Н.[1,2], Жумабаева А.[2], Дайырбаева Э.[1,4*], 2026.
[1]Институт информационных и вычислительных технологий, Алматы, Казахстан;
[2]Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан;
[3]META University, Алматы, Қазақстан;
[4]Казахский национальный исследовательский технический университет им. К.И. Сатпаева, Алматы, Казахстан.
E-mail: nurbekkyzyelmira@gmail.com

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МНОГОЯЗЫЧНЫХ ТРАНСФОРМЕРНЫХ МОДЕЛЕЙ ДЛЯ ПЕРЕВОДА С КАЗАХСКОГО ЯЗЫКА НА ГЛОССИРОВАННОЕ ПРЕДСТАВЛЕНИЕ

Турсынова Назира — инженер-программист, Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан; PhD-докторант, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: ntursynova000@gmail.com, https://orcid.org/0000-0002-1857-9028;
Еримбетова Айгерим — ведущий научный сотрудник, Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан; PhD, кандидат технических наук, ассоциированный профессор, профессор-исследователь META University, Алматы, Казахстан,
E-mail: aigerian8888@gmail.com, https://orcid.org/0000-0002-2013-1513;
Амангелді Нұрзада — старший научный сотрудник, Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан; PhD, исследователь, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: amangeldi_n_3@enu.kz, https://orcid.org/0000-0002-4669-9254;
Жумабаева Асель — PhD, старший преподаватель, Евразийский национальный университет имени Л.Н. Гумилева, Астана, Казахстан,
E-mail: zhumabayeva_as@enu.kz, https://orcid.org/0000-0002-8984-1492;
Дайырбаева Эльмира — научный сотрудник, Институт информационных и вычислительных технологий КН МНВО РК, Алматы, Казахстан; магистр, старший преподаватель, Казахский национальный исследовательский технический университет имени К.И. Сатпаева, Алматы, Казахстан,
E-mail: nurbekkyzyelmira@gmail.com, https://orcid.org/0000-0002-4255-5456.

Аннотация. *Актуальность.* Казахский язык жестов является важным направлением исследований в области инклюзивных цифровых технологий и обработки естественного языка. В последние десятилетия возрастает потребность в создании интеллектуальных систем, способных снижать барьеры коммуникации между людьми с нарушениями слуха и речи и обществом. Одной из актуальных задач является перевод текста на глоссированное представление, поскольку для казахского языка жестов отсутствуют полноразмерные параллельные корпуса, а неоднозначность жестовых единиц усложняет автоматическую обработку и синтез голосов.

Цель. Провести сравнительный анализ многоязычных трансформных моделей для перевода с казахского языка на глоссированное представление и оценить их качество, скорость работы и ресурсоемкость. Методы. В исследовании рассмотрен перевод текста в глоссы с использованием пяти моделей на основе многоязычных трансформеров: mBART-large-50, M2M100-418M, mT5-base, MarianMT и T5-base. Все модели обучались с использованием одинаковых параметров, что обеспечило сопоставимость экспериментальных результатов. Оценка качества перевода проводилась с применением распространенных метрик BLEU, ROUGE-L и chrF++, а также качественного анализа выходных последовательностей. Дополнительно анализировались вычислительная скорость моделей и уровень потребления ресурсов. Результаты и выводы. Результаты исследования показали, что более крупные многоязычные модели обеспечивают более точный перевод с казахского языка на глоссированное представление. Модели mBART-large-50 и M2M100-418M продемонстрировали более высокое качество перевода, тогда как меньшие по размеру модели оказались быстрее и менее ресурсоемкими. Полученные результаты подтверждают наличие компромисса между точностью перевода и вычислительной эффективностью. Практическая значимость исследования заключается в возможности использования сравнительного анализа при выборе архитектуры для систем автоматического перевода казахского текста в глоссы, образовательных и инклюзивных цифровых платформ, а также технологий поддержки коммуникации для людей с нарушениями слуха и речи.

Ключевые слова: казахский язык жестов, перевод текста в глоссы, глоссированное представление, многоязычные трансформеры, mBART-large-50, M2M100-418M, mT5-base, MarianMT, T5-base, BLEU, ROUGE-L, chrF++

Introduction. Kazakh sign language (KSL) is the main method of communication for deaf citizens in Kazakhstan. Most digital services still uses only Kazakh written language. Usually it is not problematic, but problem is occurred when deaf citizens have to deal with urgent information without assistance. Fact that often surprises non-professional saqwerswqerrdfts is that the sign language is not “visual version” of the spoken language: it is the separate linguistic system with own grammar. That is why providing written information (even if it uses official language) may not always mean that it is samely available for KSL users.

Practical approach to ensuring accessibility is to convert written text into intermediate symbolic layer. It is easier to deal with them then dealing with signs. In this study glossary is used as such layer. Here glossary is the sequence of the standardized tokens which corresponds to the signs. But gloss is not complete written form of KSL, rather it is the compact notation which is used for annotation and modelling. The main importance of glossary is that it allows to represent complex multimodal target to discrete sequence which simplifies usage of sequence-to-sequence methods preserving main linguistical cues. Notable that some small lexemes carry disproportionate large meaning, omitting them the meaning is changed definitely: questions may become statement or tenses may be changed. And this is not a bug but it changes the communicative intent.

Thus, translation from text to gloss can become a base for extension of technological capabilities of KSL. After development of the glossary it may be used to synthesis of signs, avatar animations, extracting sign content or editing it manually without recording new videos. Intermediate stage makes more interpretable the behavior of the system: not translation not completed successfully, output gloss may illustrate if the model lost the important phrases, incorrectly deals with negative sentences or repeatedly uses same patterns. However, progress in KSL is restricted by two structural bottlenecks. Firstly, the count of parallel corpora to translate Kazakh language to glossary is limited. Secondly, annotation process is expensive and long as it requires professional signers and consistent labeling convention: even a little difference may break control and reduce generalization.

Neural Machine Translations (NMT) provides a potential solution to these constraints through transfer learning. Of all the available methods, the Transformer architecture stands out. Transformer architecture is an encoder-decoder model based on a self-attention mechanism. Currently, Transformer is becoming very popular because it is effective in dealing with long-distance dependencies (Vaswani et al., 2017). Multilingual pretrained models provide a unified system for multiple languages. It is very well suited for languages with limited resources, such as Kazakh (Joshi et al., 2020). It can be easily applied to solve various problems by configuring it for a specific task.

The models discussed in this article, such as mBART, M2M-100, and mT5, differ in their language coverage, tokenization methods, and optimization. These solutions can help in the subsequent generations, especially when the amount of training data is limited (Liu et al., 2020; Fan et al., 2020; Xue et al., 2021). This makes them particularly attractive for Kazakh sign language, which has a small amount of parallel corpora.

However, despite the rapid development of multilingual NMT, less attention is currently being paid to sign language translation. Most of the work on sign language focuses on video-to-text translation, where the main challenges lie in recognizing images from video rather than processing text data (Camgöz et al., 2018; Camgöz et al., 2020).

On the other hand, it raises additional questions when translating from text to gloss. For example, how can existing models be adapted, to what extent can they be adjusted, and how will the size of the model affect performance?

There is also a limitation related to experimental research. In the absence of carefully prepared tests, the claimed improvement may be due to preprocessing, optimization, or decoding methods rather than the actual capabilities of a particular model.

To address the above issues, this article presents the results of a comparative analysis of five pretrained Transformer models (mBART-large-50, M2M100-418M, mT5-base, MarianMT, and T5-base). These five models were tested with similar configurations and fine tuning. They were evaluated using standard metrics such as BLEU, ROUGE-L, and chrF++. The main goal of this experimental study is to determine how multilingual pretraining performs for Kazakh sign language. The following questions were also considered: (1) Which model shows the best results in gloss generation? (2) How different are the models in terms of performance for the Kazakh language? (3) Which types of errors are most resistant to correction and why?

Literary review. Sign language is fully developed and consists of a natural language that is expressed through movements of the face, body, and hands. Sign languages have their own distinct grammatical systems and vocabularies (Bragg et al., 2019). Government agencies provide written information to deaf and hard-of-hearing people. As a result, an asymmetrical communication environment arises everywhere, as they do not have access to all the necessary government services. This is because deaf and hard-of-hearing people do not have a standardized writing system.

Therefore, sign language translation should be viewed not only as a convenience for deaf and hard-of-hearing people, but also as an important tool for improving access and reducing inequality in healthcare, education, and government digital services. Real-time sign language translation systems must meet the high latency requirements of practical applications while maintaining semantic accuracy when trained on real-world data.

These limitations are particularly important when inaccuracies in translation can have various consequences (Papatsimouli et al., 2023). From a scientific point of view, accuracy and adaptability to internal and external changes in sign language translation are considered interrelated, as they promote the use of a modeling approach rather than simple conversion.

Most modern sign language translation systems, including the SLT system, use intermediate spaces to facilitate multimodal correspondence processing (Liang et al., 2023). Using a set of methods and techniques, sign2gloss2text converts a sign language image into a text gloss, which is then translated into spoken language. Thus, a gloss can be understood as a compact text representation, since words are extracted from it and translated into sign language.

The difference between standard text form and gloss text in sign language has practical significance. Gloss sequences are subject to systematic sources of noise. In particular, annotators may differ in how they segment or mark up a given text in a given sign language. Gloss dictionaries vary across different datasets and research groups (Liang et al., 2023). However, despite their shortcomings, glosses are used continuously in practice. That is, they convert continuously moving video into a sequence of characters. This allows the use of traditional “chain-to-chain” models. Manual glossing is a costly and labor-intensive process, and there is growing interest in model architectures and training schedules that reduce the amount of costly intermediate supervision (Camgöz et al., 2020; Lin et al., 2023). Therefore, researchers are looking for an optimal way to introduce glosses that does not require a lot of labor and resources.

A lot of researchers worked closely on CSLR even before it became possible to develop truly reliable sign language translation. This is because continuous conversion of all data through recognition into translation streams based on a glossary is a necessary condition. During an experiment on continuous sign language recognition using the CSLR method, it was found that visual space and time are closely related in sign language. That is, the study found that finger movements, facial expressions indicating internal states, and head movements are interrelated. It is important to note that each of these movements conveys different grammatical information (Rastgoo et al., 2021). In the process of recognizing sign language and individual signs, the CSLR system immediately detects small changes in external gestures. In particular, changes in gestures during speech, changes in posture, as well as shading or overlapping of objects. These variations make it difficult for the system to distinguish between linguistically meaningful movements and non-linguistic movements (Cheok et al., 2019).

The researcher (Koller et al., 2018) believes that the application of deep learning has significantly improved it by systematically organizing the spatiotemporal representation of sign language data. To prove this, the Deep Sign project demonstrated that large-scale training combined with appropriate temporal modeling can lead to a significant increase in the accuracy of continuous recognition. Modern methods have been proposed to reduce the number of simultaneously perceived meaningful texts, as well as the additional gestures required to understand the meaning of these texts. However, this approach is problematic in design due to the large amount of data (Li & Meng, 2022). As a result, continuous translation work aimed at maximizing efficiency suffers in real-world conditions. That is, errors made in the perception of ambiguous texts are systematic in nature, as they have a high potential for direct propagation to subsequent translations.

Transformer-based models have been found to be highly effective in sign language translation. In particular their mechanism is so effective that it allows for modeling long distances and understanding and coordinating information from various sources, such as video and text, at a high level. The “chain-to-chain” rule is used in neurotranslation research on sign language. According to this rule, it is considered more effective to link video sequences with spoken language. Its advantage is that cascading recognition allows for systematic analysis in accordance with a single training and evaluation rule in the translation process (Camgöz et al., 2018). In response to the above-mentioned problems, Sign Language Transformers has developed a scalable structure for sign language translation. The main purpose of the device is to reduce the accumulation of errors inherent in rigid cascade pipelines by combining recognition and translation (Camgöz et al., 2018).

Some researchers have shown that glosses are not essential for the system. A method has been proposed for representing the compatibility of video and sign language translation in text without using glosses and without losing the main meaning. This method provides high-quality translation by teaching the system the relationship between sign language and speech. It uses additional signals to establish semantic links between video and text. This means that intermediate labels, which are often checked when using the model, are clearly unnecessary (Li et al., 2025). One of the main reasons for abandoning the use of glosses is to use convenient solutions and prevent data corruption (Lin et al., 2023).

On the other hand, the ability to translate sign language without the help of gloss has been proven to be poor. Yes, although it is possible to translate sign language into text using video, the second problem is that written text can only be translated with the help of gloss. To use it to convert the translated gloss into sign language (Liang et al., 2023). The use of glosses has contributed significantly to the development of sign language. The text is intended to be used to reveal meaning and is interpreted using gestures. An example of this is the Text2Sign system. The system clearly demonstrates the possibilities of creating gestures for intermediate scenes (Stoll et al., 2020). Even within a single sign language, different glossary standards are of particular interest for different language units. It is considered more appropriate to view a glossary as a program interface rather than a linguistic transcription. The conversion of text into a glossary depends not only on the architecture of the model, but also on the correct construction of the dataset. Otherwise, data truncation and other obstacles may arise (Matsumoto et al., 2006).

There are certain requirements for the gloss-to-text conversion model to work correctly. With fewer parallel corpora, the translation capabilities are more limited. This is because the quality of the translation depends on the amount of data collected. For example, suppose there is little data, but the translation of a chemical subject common to both languages into gloss will be of low quality due to the small amount of data, i.e., the parallel corpus (Haddow et al., 2022; Ranathunga et al., 2023). This requirement also applies to neural machine translation. This is because high-quality translation is not possible when working with a small amount of data in neural machine translation. (Dabre et al. , 2021).

There is also an approach to modeling based on subwords. It is used for languages with complex morphology. In this method, words taken from the dictionary are removed, and words that have not been encountered before are recognized. To increase the stability of the model, one text can be divided into several lines, and regularization can be used to visualize the quality of the model (Kudo, 2018). Another approach is to use multilingual training mode to train models on a small number of examples without any training. This feature helps translate common multilingual scenes from annotated data sources without supervision (Johnson et al., 2017). The lack of a parallel corpus for translation from Kazakh to the glossary is also exacerbated by the lack of a single standard for the glossary.

There are several approaches that have proven beneficial when translating sign language into text. First, let's look at the BART structure. It reduces noise through pre-training and effectively builds a generative model for text generation. The next model is the Transformer model, which is pre-trained using a chain-of-chains approach. This model provides reliable initialization because it encodes general linguistic patterns in its parameters (Lewis et al., 2020). mBART reduces noise similarly to the BART structure mentioned above. However, it extends pre-training to a multilingual environment, using pre-training to reduce noise. It has been shown that initializing the encoder and decoder significantly improves translation scenarios in resource-limited and domain-specific environments (Liu et al., 2020).

Text-to-text formulation in T5 allows wide range of NLP tasks including translation to be represented as conditional generation task within the same framework(Raffel et al., 2019). Further, mT5 adapted this paradigm to multilingual pretraining for more than 100 languages (Xue et al. , 2021). Simultaneously, developed a lot of other multilingual systems, such as M2M-100, which lets to deal with non-English texts (Fan et al., 2020). Also tools are essential here: Marian provides an efficient and research oriented toolkit for NMT which reduces implementation costs and evaluation proccess (Junczys-Dowmunt et al., 2018).

Overall, these findings stimulate comparative analysis in pretrained multilingual Transformers. Differences of the models are in the objectives of pretraining, in language coverage or in capacity allocation, and it may impact on performance when the training data is limited as in Kazakh Sign Language.

Kazakh Sign Language (KSL) remains considerably under-resourced in comparison with more extensively studied sign languages, and only recently has the literature begun to report reusable parallel resources. Initial efforts have focused on constructing corpora that align Kazakh written text with KSL data, thereby establishing the basic prerequisites for systematic model training and evaluation (Yerimbetova et al. , 2025). At the same time, much of the existing KSL-related research continues to concentrate on recognition-oriented tasks, such as alphabet or gesture classification. While these directions are valuable, they do not directly address the translation problem that lies at the core of many accessibility-oriented applications (Buribayev et al., 2025).

To date, there has been no controlled and reproducible comparison of multilingual encoder-decoder architectures for translation from Kazakh to gloss during training and decoding. To address this issue, this paper conducts a comparative analysis of five pre-trained Transformer models, supplemented by a targeted analysis of linguistically significant shortcomings, using a unified fine-tuning protocol and a common set of automated evaluation metrics.

In particular, there are still a number of unresolved issues in translating from Kazakh into the glossary. These include: (i) the limited and potentially inconsistent nature of glossary conventions, (ii) apparent sensitivity to tokenization choices and morphological variations, and (iii) uncertain behavior during transfer when multilingual pre-trained models are adapted to glossary notation that does not correspond to the standard written language. The situation is further complicated by evaluation: recent metric studies show that automatic evaluations may only imperfectly correlate with semantic preservation, especially for short, template-based, or appearance-dependent outputs (Lee et al., 2023).

Overall, existing research still does not provide a controlled and reproducible comparison of multilingual encoder-decoder architectures for translation from Kazakh to glossary. To fill this gap, this paper conducts a comparative analysis of five pre-trained Transformer models, supplemented by a targeted analysis of linguistically significant shortcomings, using a common tuning protocol and a common set of automatic evaluation metrics.

Materials and methods. In this section, the method of translation from Kazakh language to KSL using Transformer architecture is described. The structure of Transformer architecture is illustrated in Figure 1.

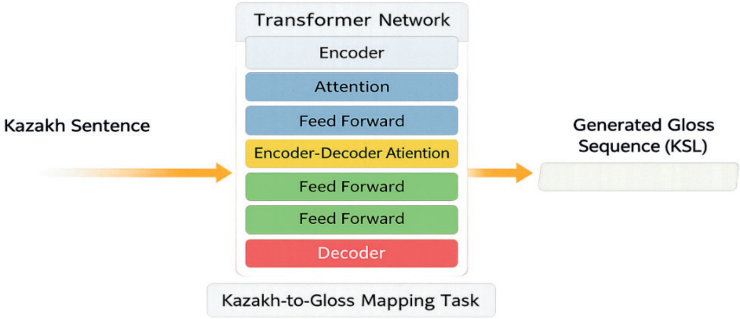

Figure 1 – Architecture of the Kazakh-to-Gloss translation model based on an encoder--decoder Transformer. The encoder processes the Kazakh input sequence using multi-head self-attention and feed-forward layers, producing contextual representations. The decoder generates the gloss sequence autoregressively using masked self-attention and encoder--decoder (cross) attention.

A. Problem Statement

Suppose that the input sentence is represented as a sequence of tokens

$$s = (s_1, s_2, \dots s_N), \tag{1}$$

and the corresponding target gloss sequence as

$$g = (g_1, g_2, \dots, g_M) \tag{2}$$

The objective is to learn a model that estimates the conditional probability of the target sequence given the source

$$P(g|s) = \prod_{t=1}^{M} P(g_t | g_{<t}\ s) \tag{3}$$

where $g_{<t}$ denotes the previously generated gloss tokens.

B. Model Architecture

Transformer architecture is an encoder–decoder model based on a self-attention mechanism. The encoder converts the input sentence into a sequence of contextual representations. The decoder outputs a sequence of tokens based on the input and output of the previous word. The internal components of the Transformer architecture are shown in Figure 1.

1) Token Embedding and Positional Encoding: To begin with, an embedding vector is created for each input toke s_i, and information about the position is stored there.

$$\mathbf{Z}^{(0)} = \mathrm{Emb}(s) + \mathrm{PosEnc}(1{:}N), \qquad (4)$$

where $\mathbf{Z}^{(0)} \in \mathbb{R}^{N\times d}$ denotes the input to the encoder stack and d is the model dimensionality. Source and target sequences are padded or truncated to a maximum length of 128 tokens.

2) Encoder Layers: The encoder consists of multiple identical layers, each comprising multi-head self-attention followed by a position-wise feed-forward network. For a given layer input , attention projections are computed as

$$\mathbf{A} = \mathbf{Z}\mathbf{W}^{(a)}, \mathbf{B} = \mathbf{Z}\mathbf{W}^{(b)}, \mathbf{C} = \mathbf{Z}\mathbf{W}^{(c)}, \qquad (5)$$

where $\mathbf{W}^{(a)}, \mathbf{W}^{(b)}, \mathbf{W}^{(c)}$ are learned projection matrices.

Scaled dot-product attention is defined as

$$\mathrm{SA}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \mathrm{softmax}\left(\frac{\mathbf{A}\mathbf{B}^{\top}}{\sqrt{d_h}}\right)\mathbf{C}, \qquad (6)$$

where d_h denotes the dimensionality of each attention head. Multi-head attention concatenates multiple such heads and applies a linear projection. Multi-head attention combines several attention mechanisms into a single representation.

Then feed-forward sublayer is applied to each word:

$$\mathrm{FF}(\mathbf{u}) = \emptyset(\mathbf{u}\mathbf{W}_1 + \beta_1)\mathbf{W}_2 + \beta_2 \qquad (7)$$

Here, $\emptyset(\cdot)$ nonlinear activation function. Residual connections and layer normalization are applied to all sublayers and then these sublayers are normalized.

3) Decoder and Cross-Attention. The decoder is similar in structure to the encoder, but it works with already generated glosses and translates them one by one. Let

$$\mathbf{R}^{(0)} = Emb(\mathbf{g}_{<t}) + \mathrm{PosEnc}(1{:}t-1) \qquad (8)$$

be the decoder input representation. Masked self-attention applies a causal mask M:

$$\mathrm{MSA} = \mathrm{softmax}\left(\frac{\mathbf{AB}^{\intercal} + \mathrm{M}}{\sqrt{d_h}}\right)\mathbf{C} \qquad (9)$$

Encoder--decoder attention conditions the decoder on encoder outputs E**:**

$$\mathrm{CA}(\mathbf{D},\mathbf{E}) = \mathrm{softmax}\left(\frac{\mathbf{DE}^{\intercal}}{\sqrt{d_h}}\right)\mathbf{E} \qquad (10)$$

allowing the model to align gloss tokens with relevant source representations.

4) Output Projection: The decoder state r_t is mapped to the vocabulary distribution as

$$P(g_t|\mathbf{g}_{<t}\ \mathrm{s}) = \mathrm{softmax}(\mathbf{W}_o\mathbf{r}_t + \gamma), \qquad (11)$$

where W_0 and γ are learned parameters.

C. Optimization Objective

During training, the model compares predicted gloss with correct one and is penalized for each mismatch. As a result, the model learns t bring its predictions closer to the correct translation:

$$\mathcal{J} = -\sum_{t-1}^{M} \log P(g_t|\mathbf{g}_{<t}\ \mathrm{s}) \qquad (12)$$

Empty tokens (padding) are excluded from the process so they don’t affect the training.

D. Training Configuration

All models are trained under the same conditions with same configurations. Hugging Face Transformers library is used. Experiments are conducted on an NVIDIA RTX 3080 GPU with 10 GB of memory. The learning rate is set to 3×10^{-5}, the batch size is 4, and each model is trained for five epochs. Mixed-precision training is employed in order to reduce memory consumption and improve computational efficiency.

E. Evaluation

Model performance is evaluated using BLEU, computed with SacreBLEU. The BLEU score is defined as

$$\mathrm{BLEU} = \mathrm{BP} \cdot \exp\left(\sum_{n=1}^{N} \alpha_n \log \pi_n\right), \tag{13}$$

where π_n denotes modified n -gram precision, α_n are uniform weights, and BP is the brevity penalty. After each epoch, the model is evaluated on test sets.

Results. This section reports both quantitative and qualitative results for the Kazakh to gloss translation task. Five Transformer-based models are evaluated under identical training conditions, and their performance is compared in terms of automatic evaluation metrics, training efficiency, and representative translation examples.

A. Overall Performance Comparison

Figure 2 presents a comparison of BLEU scores across the evaluated Transformer models, together with their corresponding training times. Among all models, mBART-large-50 achieves the highest BLEU score, followed closely by M2M100-418M. Models with fewer parameters, such as MarianMT and T5-base, exhibit lower BLEU scores but substantially reduced training time.

The results highlight a clear trade-off between translation quality and computational cost. Larger multilingual pretrained models benefit from richer cross-lingual representations, which translate into higher accuracy for gloss generation.

B. Quantitative Evaluation Metrics

Table 1 summarizes the BLEU, ROUGE-L, and chrF++ scores obtained by all evaluated models. The relative ordering of systems is consistent across metrics, with mBART-large-50 achieving the best overall performance, followed by M2M100-418M and mT5-base.

The nearly seven-point BLEU difference between mBART-large-50 and T5-base suggests that both model capacity and multilingual pretraining substantially influence the preservation of semantic content in gloss generation. ROUGE-L and chrF++ exhibit comparable patterns, indicating that higher BLEU scores are accompanied by improvements in both sequence-level alignment and character-level accuracy.

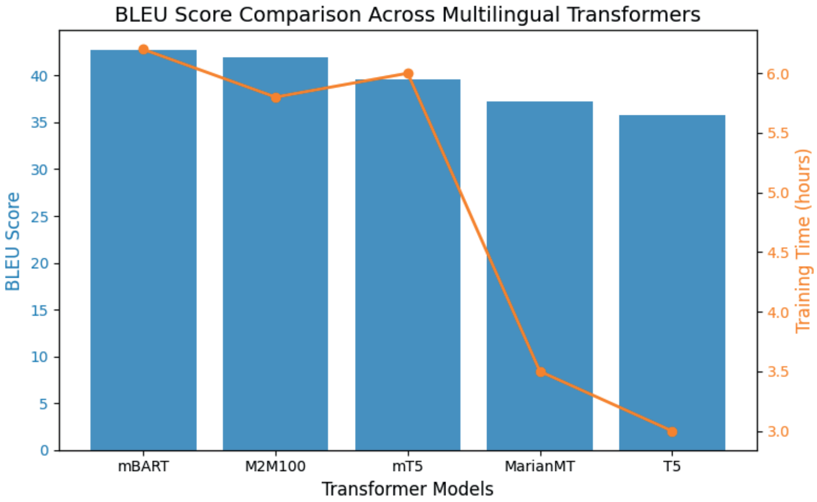

Figure 2 - Comparison of BLEU scores and training time across Transformer models for Kazakh-to-Gloss translation. Larger multilingual models achieve higher BLEU scores at the cost of increased training time.

Table 1 – Automatic evaluation results for Kazakh-to-Gloss translation

Model	BLEU	ROUGE-L	chrF++
mBART-large-50	42. 7	0. 78	0. 61
M2M100-418M	41. 9	0. 77	0. 60
mT5-base	39. 6	0. 75	0. 58
MarianMT	37. 2	0. 73	0. 55
T5-base	35. 8	0. 70	0. 53

scores correspond to improved sequence-level alignment and character-level accuracy.

C. Model Size and Training Efficiency

Table 2 reports the number of parameters and the observed training time for each model. While mBART-large-50 and mT5-base have similar parameter scales, mT5-base attains slightly lower BLEU scores, suggesting that differences in architecture and pretraining objectives influence downstream performance beyond raw model size alone.

Table 2 – Model size and training time comparison

Model	Parameters (M)	Training Time (hrs)
mBART-large-50	610	6. 2
M2M100-418M	418	5. 8
mT5-base	580	6. 0
MarianMT	300	3. 5
T5-base	220	3. 0

MarianMT and T5-base require roughly half the training time of mBART-large-50, but incur a reduction of approximately 5–7 BLEU points. Such a trade-off may be acceptable in settings where computational resources are constrained or rapid experimentation is prioritized.

D. Qualitative Analysis

Table 3 provides representative translation examples that illustrate common error patterns. Higher-performing models, including mBART and M2M100, tend to preserve temporal expressions, interrogative markers, and canonical gloss word order. By contrast, smaller models more frequently omit modifiers or question particles and show greater variability in word ordering.

Table 3 – Qualitative examples of Kazakh-to-Gloss generation with per-example BLEU (in parentheses)

Kazakh inpu	Reference gloss	mBART	M2M100	MarianMT	T5
Мен үйге ерте келемін.	МЕН ҮЙ ЕРТЕ КЕЛ	МЕН ҮЙ ЕРТЕ КЕЛ (100)	МЕН ҮЙ ЕРТЕ КЕЛ (100)	МЕН ҮЙ КЕЛ ЕРТЕ (54)	МЕН ҮЙ КЕЛ (41)
Біз дүкеннен нан сатып алдық.	БІЗ ДҮКЕН НАН САТЫП АЛ	БІЗ ДҮКЕН НАН САТЫП АЛ (100)	БІЗ ДҮКЕНДЕН НАН САТЫП АЛ(82)	БІЗ НАН АЛ ДҮКЕН (46)	БІЗ НАН АЛ (33)
Сен ертең университетке барасың ба?	СЕН ЕРТЕҢ УНИВЕРСИТЕТ БАР МА	СЕН ЕРТЕҢ УНИВЕРСИТЕТ БАР МА (100)	СЕН ЕРТЕҢ УНИВЕРСИТЕТКЕ БАР МА (78)	СЕН ЕРТЕҢ УНИВЕРСИТЕТ БАР (52)	СЕН УНИВЕРСИТЕТ БАР (38)

Discussions. These examples indicate that interrogative particles and temporal modifiers are especially prone to omission in lower-capacity models. Such errors are of particular importance in sign language contexts, where these elements are often realized through distinct facial expressions or specific signing patterns.

Overall, the experimental findings reveal a consistent relationship between model capacity, multilingual pretraining, and Kazakh-to-gloss translation quality. Larger models, including mBART-large-50 and M2M100, achieve higher BLEU, ROUGE-L, and chrF++ scores, reflecting improved preservation of both semantic content and gloss-level structure. This performance gain, however, is accompanied by increased training time and computational cost.

By contrast, smaller models such as MarianMT and T5-base converge more quickly and require substantially fewer resources, but they exhibit systematic omissions of temporal modifiers and interrogative markers in qualitative analyses. These shortcomings are particularly consequential for sign language processing, where such elements frequently correspond to grammatical distinctions and non-manual features.

The resulting trade-off between accuracy and efficiency suggests that model selection should be informed by application-specific constraints. High-precision gloss generation pipelines are better served by larger multilingual models, whereas resource-limited scenarios may favor smaller architectures despite reduced semantic completeness.

Taken together, the reported results establish strong baseline performance for Kazakh-to-gloss translation and offer empirical guidance for the development of future Kazakh Sign Language systems, including downstream sign synthesis and avatar-based generation.

Conclusion. This paper presented a systematic investigation of Transformer-based models for Kazakh-to-gloss translation as a foundational step toward computational processing of Kazakh Sign Language. By formulating the task within a sequence-to-sequence framework and evaluating multiple multilingual pretrained architectures under a unified experimental protocol, we established clear performance baselines for this under-resourced setting.

The experimental results show that larger multilingual models, most notably mBART-large-50 and M2M100-418M, deliver superior performance across BLEU, ROUGE-L, and chrF++ metrics. In contrast, smaller models achieve shorter training times but exhibit reduced semantic completeness. Qualitative analyses further indicate that lower-capacity models tend to omit temporal modifiers and interrogative markers, elements that play an important linguistic role in sign language representations.

Taken together, these findings underscore a practical trade-off between translation quality and computational efficiency, suggesting that model choice should be informed by the requirements of the intended application. This study provides an initial benchmark for Kazakh-to-gloss translation and establishes a foundation for future work on end-to-end Kazakh Sign Language systems, including integration with sign synthesis components and evaluation involving native signers.

References

Bragg D., Koller O., Bellard M., Berke L., Boudreault P., Braffort A., Caselli, N., Huenerfauth M., Kacorri H., Verhoef T., Vogler C., & Morris M.R. (2019) Sign language recognition, generation, and translation: An interdisciplinary perspective. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*. — P. 16–31. https://doi.org/10.1145/3308561.3353774 (in Eng.)

Buribayev Z., Aouani M., Zhangabay Z., Yerkos A., Abdirazak Z., & Zhassuzak M. (2025) Enhancing Kazakh Sign Language recognition with BiLSTM using YOLO keypoints and optical flow. *Applied Sciences, 15*(10). — 5685 p. https://doi.org/10.3390/app15105685 (in Eng.)

Camgöz N. C., Hadfield S., Koller O., Ney H., & Bowden R. (2018) Neural sign language translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. — P. 7784–7793. https://doi.org/10.1109/CVPR. 2018.00812 (in Eng.)

Camgöz N.C., Koller O., Hadfield S., & Bowden R. (2020) Sign language transformers: Joint end-to-end sign language recognition and translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. — P. 10020–10030. https://doi.org/10.48550/arXiv.2003.13830 (in Eng.)

Cheok M.J., Omar Z., & Jaward M.H. (2019) A review of hand gesture and sign language recognition techniques. *International Journal of Machine Learning and Cybernetics, 10*. — P. 131–

153. https://doi.org/10.1007/s13042-017-0705-5 (in Eng.)

Dabre R., Chu C. , & Kunchukuttan A. (2021) A survey of multilingual neural machine translation. *ACM Computing Surveys, 53*(5), Article 99. https://doi.org/10.1145/3406095 (in Eng.)

Fan A., Bhosale S., Schwenk H., Ma Z., El-Kishky A., Goyal S., Baines M., Celebi O., Wenzek G., Chaudhary V., Goyal N. Birch T., Liptchinsky V., Edunov S., Grave E., Auli M., & Joulin A. (2020) Beyond English-centric multilingual machine translation. *arXiv*. https://doi.org/10.48550/arXiv.2010.11125 (in Eng.)

Haddow, B., Kocmi, T., Birch, A., & Bogoychev, N. (2022). Survey of low-resource machine translation. *Computational Linguistics, 48*(3). — P. 673–732. https://doi. org/10. 1162/coli_a_00446 (in Eng.)

Johnson M., Schuster M., Le Q.V., Krikun M., Wu Y., Chen Z., Thorat N., Viégas F., Wattenberg M., Corrado G., Hughes M., & Dean J. (2017) Google's multilingual neural machine translation system: Enabling zero-shot translation. *Transactions of the Association for Computational Linguistics, 5*. — P. 339–351. https://doi.org/10.1162/tacla00065 (in Eng.)

Koller O., Zargaran O., Ney H., & Bowden R. (2018) Deep sign: Enabling robust statistical continuous sign language recognition via deep learning. *International Journal of Computer Vision, 126*. — P. 1311–1325 https://doi.org/10.1007/s11263-018-1121-3 (in Eng.)

Kudo T. (2018) Subword regularization: Improving neural network translation models with multiple subword segmentations. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL)*. — P. 66–75. https://doi.org/10.48550/arXiv.1804.10959(in Eng.)

Lee J., Kim H., & Lee J. (2023) A survey of automatic metrics for machine translation evaluation. *Mathematics, 11*(4), 1006. https://doi.org/10.3390/math11041006 (in Eng.)

Lewis M., Liu Y., Goyal N., Ghazvininejad M., Mohamed A., Levy O., Stoyanov V., & Zettlemoyer L. (2020) BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. — P. 7871–7880. https://doi.org/10.18653/v1/2020.acl-main.703 (in Eng.)

Li H., & Meng J. (2022) Multi-view sign language recognition with deep learning: A review and new framework. *Applied Intelligence, 52*. — P. 14534–14560 https://doi.org/10.1007/s10489-022-03407-5

Li X., Zhang Y., & Wang Z. (2025) Gloss-free sign language translation based on fusion attention. *Applied Soft Computing, 185*, 113848. https://doi.org/10.1016/j.asoc.2025.113848 (in Eng.)

Liang Z., Li H., & Chai J. (2023) Sign language translation: A survey of approaches and techniques. *Electronics, 12*(12), 2678. https://doi. org/10. 3390/electronics12122678 (in Eng.)

Lin K., Wang X., Zhu L., Sun K., Zhang B., & Yang Y. (2023) Gloss-free end-to-end sign language translation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL)*. — P. 12904–12916. https://doi.org/10.48550/arXiv.2305.12876 (in Eng.)

Liu Y., Gu J., Goyal N., Li X. , Edunov S. , Ghazvininejad M. , Lewis M. , & Zettlemoyer L. (2020) Multilingual denoising pre-training for neural machine translation. *arXiv*. https://doi.org/10.48550/arXiv.2001.08210 (in Eng.)

Papatsimouli M., Sarigiannidis P., & Fragulis G. (2023) A survey of advancements in real-time sign language translators: Integration with IoT technology. *Technologies, 11*(4), 83. https://doi.org/10. 3390/technologies11040083 (in Eng.)

Raffel C., Shazeer N., Roberts A., Lee K., Narang S., Matena M., Zhou Y., Li W., & Liu P.J. (2019) Exploring the limits of transfer learning with a unified text-to-text transformer. *arXiv*. https://doi.org/10.48550/arXiv.1910.10683 (in Eng.)

Ranathunga S., Plank B., & El-Kishky A. (2023) Neural machine translation for low-resource languages: A survey. *ACM Transactions on Asian and Low-Resource Language Information Processing*. https://doi. org/10. 1145/3567592 (in Eng.)

Rastgoo R., Kiani K., & Escalera S. (2021) Sign language recognition: A deep survey. *Expert Systems with Applications, 164*, 113794. https://doi.org/10.1016/j.eswa.2020.113794 (in Eng.)

Stoll S., Camgöz N., Hadfield S., & Bowden R. (2020) Text2Sign: Towards sign language production using neural machine translation and generative adversarial networks. *International Journal of Computer Vision, 128*. — P. 891–908. https://doi.org/10.1007/s11263-019-01281-2 (in Eng.)

Xue L., Constant N., Roberts A., Kale M., Al-Rfou R., Siddhant A., Barua A., & Raffel C. (2021) mT5: A massively multilingual pre-trained text-to-text transformer. *arXiv*. https://doi.org/10.48550/arXiv.2010.11934 (in Eng.)

Yerimbetova A., Sakenov B., Sambetbayeva M., Daiyrbayeva E., Berzhanova U., & Othman M. (2025) Creating a Parallel Corpus for the Kazakh Sign Language and Learning. *Applied Sciences*, *15*(5), 2808. https://doi.org/10.3390/app15052808 (in Eng.)

Publication Ethics and Publication Malpractice in the journals of the Central Asian Academic Research Center LLP

For information on Ethics in publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the journals of the Central Asian Academic Research Center LLP implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The Central Asian Academic Research Center LLP follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the Central Asian Academic Research Center LLP.

The Editorial Board of the Central Asian Academic Research Center LLP will monitor and safeguard publishing ethics.

Requirements for articles design for publication in the journal are available on the websites:
www:nauka-nanrk.kz
http://physics-mathematics.kz/index.php/en/archive
ISSN2518-1726 (Online),
ISSN 1991-346X (Print)

Managing Editor: *A.Shormakova*
Editors: *D.S. Alenov, T. Apendiev*
Computer layout: *G.D. Zhadyranova*

Signed for print: June 15, 2026
Format: 70×90 1/16. 26.5 printed sheets. Order No. 2.

«Central Asian Academic Research Center» LLP
Almaty, Shevchenko Street, 28