applied sciences

Article

Creating a Parallel Corpus for the Kazakh Sign Language and Learning

Aigerim Yerimbetova [1,2], Bakzhan Sakenov [1], Madina Sambetbayeva [1,3], Elmira Daiyrbayeva [1,4], Ulmeken Berzhanova [1,5,*] and Mohamed Othman [6]

1 Laboratory of Computer Engineering of Intelligent Systems, Institute of Information and Computational Technologies, Almaty 050010, Kazakhstan; aigerian8888@gmail.com or aigerimy@illinois.edu (A.Y.); sbakzhan22@gmail.com (B.S.); madina_jgtu@mail.ru (M.S.); nurbekkyzy_e@mail.ru (E.D.)
2 Global Education and Training, University of Illinois Urbana-Champaign, Champaign, IL 61820, USA
3 Department of Information Systems, L.N. Gumilyov Eurasian National University, Astana 010008, Kazakhstan
4 Department of Software Engineering, Satbayev University, Almaty 050010, Kazakhstan
5 Faculty of Information Technology, Al-Farabi Kazakh National University, Almaty 050038, Kazakhstan
6 Department of Communication Technology and Network, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang 43400, Malaysia; mothmanupm@gmail.com
* Correspondence: berzhanovaulmekenn@gmail.com

Abstract: Kazakh Sign Language (KSL) is a crucial communication tool for individuals with hearing and speech impairments. Deep learning, particularly Transformer models, offers a promising approach to improving accessibility in education and communication. This study analyzes the syntactic structure of KSL, identifying its unique grammatical features and deviations from spoken Kazakh. A custom parser was developed to convert Kazakh text into KSL glosses, enabling the creation of a large-scale parallel corpus. Using this resource, a Transformer-based machine translation model was trained, achieving high translation accuracy and demonstrating the feasibility of this approach for enhancing communication accessibility. The research highlights key challenges in sign language processing, such as the limited availability of annotated data. Future work directions include the integration of video data and the adoption of more comprehensive evaluation metrics. This paper presents a methodology for constructing a parallel corpus through gloss annotations, contributing to advancements in sign language translation technology.

Keywords: Kazakh sign language; parallel corpus; sign language; machine translation; deep learning; sequence to sequence model

Academic Editor: Douglas O'Shaughnessy

Received: 12 December 2024
Revised: 23 February 2025
Accepted: 26 February 2025
Published: 5 March 2025

Citation: Yerimbetova, A.; Sakenov, B.; Sambetbayeva, M.; Daiyrbayeva, E.; Berzhanova, U.; Othman, M. Creating a Parallel Corpus for the Kazakh Sign Language and Learning. *Appl. Sci.* **2025**, *15*, 2808. https://doi.org/10.3390/app15052808

1. Introduction

According to the World Health Organization (WHO), the global population of deaf individuals reached approximately 466 million in 2023, with projections indicating an increase to 900 million by 2050 [1]. "Hearing loss" is a broad term encompassing conditions ranging from mild to severe impairments. For many affected individuals, sign language serves as their primary mode of communication. However, they often face challenges in interacting with the broader society. In Kazakhstan, according to data from the Ministry of Labor and Social Protection, the number of officially recognized individuals with disabilities stood at 715,157 as of May 2022 [2].

Sign language is a visual communication system that relies on hand gestures, facial expressions, and body movements to convey information. It is widely used by deaf and hard-of-hearing individuals, enabling them to communicate with each other and, in some cases, with hearing individuals who do not understand sign language [3]. Most machine translation systems primarily focus on spoken languages, leaving translation between spoken and signed languages significantly underdeveloped. Unlike spoken languages, sign language is a visual–spatial system that conveys meaning through a combination of hand movements, finger positions, facial expressions, and body postures in three-dimensional space. This fundamental difference makes the process of translating written text into sign language particularly challenging, as there is no universally standardized grammatical structure across sign languages. Despite these challenges, various strategies exist for converting text into sign language, where the written text serves as the source and the documented sign language representation as the output. Studies [4–10] explore various approaches to sign language machine translation, including rule-based methods, machine learning techniques, avatar-based translation, and the development of sign language corpora. A key challenge in translating written language into sign language is the lack of uniform grammatical structures, which complicates the process and limits translation accuracy. However, these methods play a crucial role in improving communication accessibility, allowing sign language users to engage more effectively in today's information-driven society.

The development and training of a parallel corpus for KSL is a crucial research area that significantly enhances communication and fosters the integration of individuals with hearing and speech impairments into Kazakh society. Sign language serves as a primary mode of communication for those with limited ability to hear and produce spoken language. In Kazakhstan, as in many other countries, individuals with hearing and speech impairments face challenges in education, access to information, and social integration. These difficulties stem from the limited awareness and dissemination of knowledge about KSL, as well as the lack of research and resources dedicated to its study and development.

Research initiatives in KSL aim to address these challenges by ensuring that deaf and deafblind individuals have access to education and information while deepening the understanding of their cultural and linguistic backgrounds. The development of a parallel KSL corpus and the advancement of KSL instruction have the potential to serve as valuable tools for researchers, educators, and professionals working with deafblind individuals. These efforts contribute to the creation of new KSL teaching methodologies, the development of translation resources, and the expansion of linguistic knowledge of sign language. This section of the paper reviews the key motivations and challenges faced by researchers in the fields of KSL and parallel corpus development.

The key challenges faced by researchers in the field of KSL and parallel corpus development include the following:

- *Preservation and Dissemination of Cultural Heritage:* KSL is not only a means of communication but also an essential part of the cultural heritage of individuals with hearing and speech impairments in Kazakhstan. The development of a parallel corpus plays a crucial role in preserving and promoting this cultural asset, ensuring its continued use and recognition.
- *Promoting Access to Education:* Individuals with hearing and speech impairments often struggle to access education due to the limited distribution of KSL. A parallel corpus can support the development of KSL teaching materials and methodologies, making education more accessible.
- *Facilitating Communication and Collaboration*: The creation and training of a KSL corpus enable the development of tools that improve communication between deaf and hearing individuals. These tools can facilitate translation from KSL into text or spoken language, thereby enhancing accessibility.

- *Advancing Linguistic Research:* A parallel corpus serves as a valuable resource for studying KSL's grammatical structures, vocabulary, and linguistic organization, contributing to advancements in sign language linguistics.
- *Promoting Social Integration and Inclusion:* Strengthening KSL and developing a well-structured corpus contribute to the social integration of individuals with hearing and speech impairments, allowing them to participate more actively in society.
- *Supporting Professional Interpreters:* A parallel corpus provides an essential resource for interpreters and translators working with deaf and hard-of-hearing individuals, enabling more accurate and effective communication.

These challenges and goals drive researchers in the field of KSL and parallel corpus development, fostering efforts to enhance communication, integration, and quality of life for people with hearing and speech impairments in Kazakhstan and beyond.

A key achievement of this project is the creation of a parallel corpus of Kazakh-language texts and KSL glosses, which greatly advances linguistic research and supports language acquisition for KSL users. This study addresses the limited availability of annotated sign language data, enabling statistical analysis and the development of machine learning models that improve in accuracy as data volume increases. This initiative is particularly groundbreaking in Kazakhstan, where such resources have previously been unavailable.

The development of a custom parser to automate the transformation of Kazakh-language text into KSL glosses represents a significant technical innovation. This parser accelerates data generation for machine translation models while mitigating the limitations of manual annotation. Using a rule-based approach informed by an analysis of KSL word order and grammar, it effectively processes unique syntactic structures, ensuring precise alignment between Kazakh text and KSL glosses.

This research presents an in-depth linguistic analysis of KSL grammar, focusing on key features such as word order, animacy, reversibility, and locative cases. This study highlights significant differences between spoken Kazakh and KSL, providing valuable insights into KSL's grammatical structure. These findings directly inform the design of the parser and the architecture of machine translation models, ensuring alignment with KSL's linguistic properties.

The application of Transformer-based deep learning models for Kazakh-to-KSL gloss translation represents a novel contribution. This study demonstrates the effectiveness of Transformers in capturing complex dependencies within KSL, achieving high translation accuracy as validated by BLEU scores. Their ability to process long sequences and model spatial–temporal relationships makes them particularly well-suited for the intricacies of sign language translation.

The dataset and methodologies developed in this study contribute to significant advancements in gesture recognition. By employing techniques such as tokenization, morphological analysis, and lemmatization, the model extracts meaningful linguistic features, enhancing the accuracy of Kazakh-to-KSL translation. This dataset supports the application of deep learning algorithms, particularly Transformers, which surpass traditional methods by offering high recognition accuracy, scalability for large datasets, and adaptability to diverse contexts.

This research highlights challenges related to the limited availability of annotated data and the diversity of gestures in KSL. Addressing these obstacles through comprehensive annotations and expanded data sources has the potential to significantly advance sign language recognition technologies. Such innovations will improve access to education and information for individuals with hearing and speech impairments, fostering greater inclusivity and technological progress.

The article includes the following sections:

- A review of the existing research in sign languages (SL).
- An explanation of the methodology for recording sign language symbols.
- An analysis of sentence structure in KSL.
- A discussion of machine translation principles using a Kazakh glossary.
- A description of the parameters and implementation details of the proposed method.

2. Related Works

A parallel corpus plays a crucial role in the development of avatar-based sign language translation systems. As part of this study [11], a comprehensive review of the existing research was conducted. One study focused on American Sign Language (ASL) and introduced a new algorithm for converting English sentences into ASL. The researchers developed a parser that automatically translates sentences and includes an interface for user interaction. The study utilized a dataset of 880 words, including English glosses and ASL, combined with transformation rules to construct a bilingual text corpus containing 800 million words. However, it is important to note that this corpus did not account for semantic aspects or verb types in sign language, such as "agreement" and "disagreement" verbs.

In a related study [12], a German public television channel, PHOENIX, broadcast news and weather forecasts with sign language interpretation daily over a three-year period (2009–2011). Transcription was conducted for weather forecasts from 386 news broadcasts, performed by native deaf and hard-of-hearing speakers of German Sign Language (DGS). Additionally, the German weather forecasts were transcribed semi-automatically using the RASR speech recognition system. To enhance translation possibilities, an additional version of the glossaries was translated into spoken German.

In 2018, the DGS-Korpus project team released the first publicly accessible version of the DGS corpus, marking a significant milestone in its development [13]. Initially, the focus was on expanding the corpus, but following its release, efforts shifted toward content enrichment. New data formats were introduced, annotation agreements were established, and OpenPose data for all transcripts were published. The research portal and community websites continued to evolve, incorporating persistent identifiers, archival versions of past releases, and mobile-friendly enhancements. The portal itself underwent substantial upgrades, including an improved web transcript viewer, a KWIC comparison tool, links to other DGS linguistic resources, and a German-language interface in addition to English. This article outlines the evolution of the DGS corpus, covering its first release in 2018, second release in 2019, and third release in 2020.

Several key methods have been employed in the field of sign language translation research [14–16], each aiming to enhance translation performance and improve model efficiency across diverse datasets. (a) The first study introduced a key-point normalization method to improve the direct translation of sign language videos into spoken language. This approach uses stochastic frame selection and enlargement techniques to adapt to different body parts and enhance translation accuracy. (b) The second study proposed an encoder-decoder-based Transformer architecture, enabling continuous sign language recognition and translation in parallel. This model simplifies data collection and improves sign language recognition without gloss-level annotations, achieving a 30% efficiency increase. (c) The third study developed an alternative key-point normalization method for sign language translation without gloss annotations. The model was adapted to multiple databases, reducing information loss in videos through an improved frame selection technique, with its effectiveness confirmed through experimental evaluations.

These methods collectively contribute to advancements in sign language recognition and translation by improving data processing efficiency and enhancing model performance, particularly when applied across diverse datasets.

Additionally, research in lexicography and parallel corpus development [17,18] has explored the composition, structure, and effectiveness of sign language corpora. Two key studies were analyzed: the Anglo hryvian parallel Corps (EGPC) and the Swiss three Sign Language parallel Corps (SwissSLi). Both studies introduce new opportunities for lexicography and machine translation. However, each face unique challenges. The EGPC study is hindered by the scarcity of resources for the Georgian language, while SwissSLi struggles with ensuring the quality and completeness of parallel data. Despite these obstacles, both projects hold significant potential for driving progress in lexicographic research and machine translation. However, further research and an expanded resource base are essential to maximize their impact.

3. Materials and Methods

3.1. Glossing Signs

Reference [19] presents one of the earliest annotation systems for interpreting sign language. Until then, gestures were considered indivisible units with no internal structure. The Stokoe system for American Sign Language (ASL), which utilized graphic symbols, was among the first methods for recording these elements. Later, additional systems were developed, including HamNoSys [20] and the gloss recording system [21].

Glossing is a method of annotating sign language by assigning a written form to gestures. While it is not a direct translation, it serves as a structured approximation. Glosses do not provide an exact translation but instead follow the sequential word order of Kazakh, aligning with its equivalent sign language structures. This practice has evolved in linguistics to aid the study and analysis of languages, particularly those without a written form. In this work, we use glosses to describe KSL. The next section presents step-by-step instructions for creating our parallel corpus.

Our research utilizes modern software tools such as Python 3.10, Stanza 1.10.1, PyTorch 2.1, NumPy 2.2.0, and Pandas 2.2.3. Stanza is used for natural language processing tasks, including morphological and syntactic analysis. PyTorch enables efficient training of deep neural networks, while NumPy and Pandas facilitate working with multidimensional data arrays and tabular structures, streamlining data preprocessing and result analysis.

3.2. Word Order in KSL

The arrangement of words plays a crucial role in the grammar of any spoken language. Unlike spoken languages, where words are pronounced sequentially due to the constraints of the speech apparatus, sign languages differ in this aspect. They are not entirely sequential, as both hands can be used simultaneously to produce different symbols, reducing the need for strict word order. In this study, we cannot definitively claim that sentence structure in sign languages serves the same grammatical function as in spoken languages. This question is central to our research. Given the importance of sentence structure, it is necessary to examine whether it functions as a grammatical tool in both signed and spoken languages or whether sign languages exhibit unique structural characteristics due to their visual–spatial nature. Based on these questions, the aim of this study is to analyze the distribution of key sentence elements (subject, object, and verb) in simple declarative sentences in KSL and explore the possibility of identifying its basic sentence structure. This study also emphasizes the importance of a reliable methodology. KSL is used by individuals with hearing and speech impairments in Kazakhstan and some other countries of the former Soviet Union, including Russia and Belarus. Until recently, linguistic research on Russian Sign Language (RSL) was virtually nonexistent, with only a few studies emerging in recent years [22]. The word order in KSL has not yet been systematically examined, but [23] suggests that it is flexible. Word order is a phenomenon that can be observed without specific constraints. However, determining what constitutes "basic word order" and whether a language has a default word order is a more complex issue. Before delving into KSL, we will first discuss the concept of basic word order in both spoken and sign languages. Additionally, we will outline the methodology used for collecting and analyzing RSL data. Following this, we will present the findings of our research. Finally, we will conclude with a discussion on the issue of basic word order in KSL.

3.3. Analysis of the Word Order in the KSL

As previously noted, one of the main challenges in analyzing KSL is the lack of available data, particularly in glossary format. Written representations of signs reflect spoken language, making them easier to interpret. Figure 1 presents the syntactic structure of a Kazakh sentence and its corresponding transformation into sign language (See Figure 1). The upper layer of the diagram represents the Kazakh sentence, broken down into its syntactic components, including: ADV (adverb, "КЕШЕ" (YESTERDAY)), NP (noun phrase, including ANT/M—determiner, "АСЛАН" (ASLAN), and NOM—noun), CNJCOO (conjunction), TV (transitive verb, "САТЫП" (BUY)), VAUX (auxiliary verb, "АЛДЫ" (DID)), and punctuation (question mark).

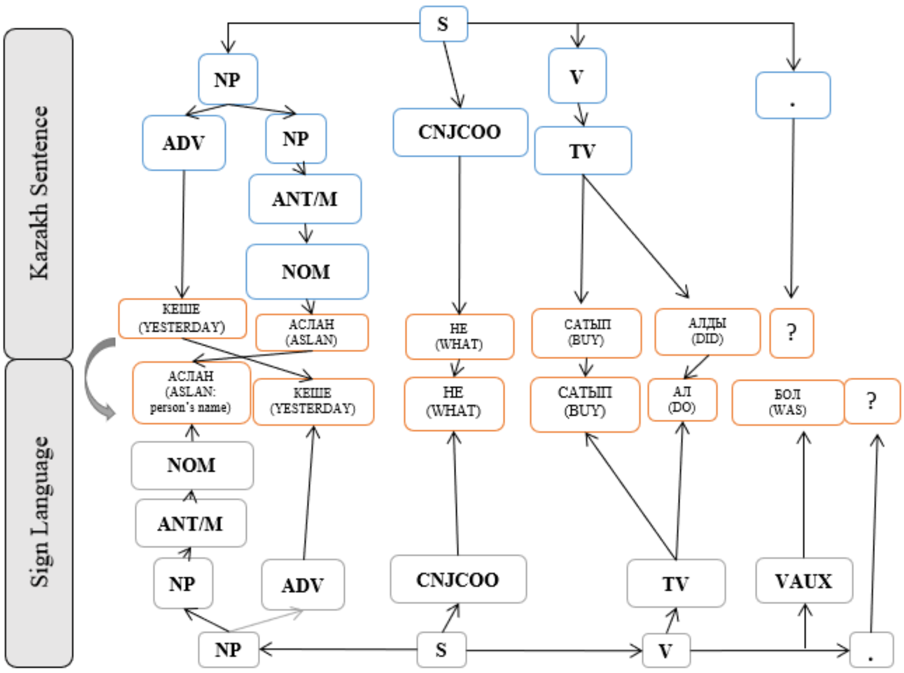


Figure 1. An example of the grammatical structure of the KSL.

The lower layer represents the corresponding components in sign language, where the structure is adapted to account for syntactic differences.

The relationships between these layers are illustrated with arrows, indicating changes in word order. For instance, the adverb "КЕШЕ" (YESTERDAY) is repositioned from the beginning of the sentence to a later position. Similarly, word order is adjusted to align with sign language grammar, where the subject (e.g., "АСЛАН" (ASLAN)) may appear after circumstances or objects.

The diagram visually demonstrates how a Kazakh sentence is transformed into its sign language equivalent, highlighting key syntactic differences and adaptations necessary for accurate representation in KSL.

The word order in sign languages is influenced by morphosyntactic, semantic, pragmatic, and modality factors [24]. Basic word orders such as subject-verb-object (SVO) and subject-object-verb (SOV) emerge under specific conditions, including topicalization, verb classification, classifier usage, and aspectual marking. In some languages, simple verbs follow the SVO order, while consistent verbs adhere to SOV. Constructions involving classifiers and aspectually marked verbs often deviate from the basic word order due to increased morphological complexity and aspectual labeling. Factors such as frequency, distribution, and simplicity of word order contribute to determining the fundamental sentence structure in a language.

For this study, more than 500 natural language sentences were analyzed and categorized into seven groups: reversible, nonreversible, locative, animate, inanimate, heavy, and nonheavy.

Reversible–Nonreversible: As observed in other sign languages, reversibility affects word order in KSL. In reversible situations, the preferred word order is SVO, while in nonreversible contexts, SOV is more commonly used.

Locative: The most frequent word order in locative sentences is OSV, though SOV and OVS are also possible in many sign languages. In locative constructions, the location is generally established first, followed by the positioning of the object.

Animate–Inanimate: Sentences with animate objects tend to favor the VO order, whereas inanimate objects typically follow an OV pre-verbal order.

Heavy–Nonheavy: In this classification, "heavy" refers to gestures that require considerable physical effort, such as large movements or forceful execution. "Nonheavy" gestures, in contrast, demand less physical effort and involve smaller movements.

An analysis of 163 sentences (See Table 1), categorized based on reversibility, revealed the following patterns:

Table 1. Sentence Order Distribution in Reversible and Nonreversible Situations.

Topic		Count Sentence	%
Reversible	SOV	29	44.61
	SVO	35	53.85
	OSV	1	1.54
	Total	65	100
Nonreversible	SOV	37	37.75
	SVO	44	44.90
	VO	17	17.35
	Total	98	100

In reversible situations, the predominant order is SVO, occurring in 35 sentences (53.85%). The SOV order appears in 29 sentences (44.61%), while OSV is found in only 1 sentence (1.54%). This indicates a strong preference for the SVO structure in reversible contexts.

In nonreversible situations, SVO remains the most frequent order, observed in 44 sentences (44.90%). The SOV order is present in 37 sentences (37.75%), while the VO (verb-object) structure is found in 17 sentences (17.35%).

Thus, the SVO order is dominant in both categories, though it is more pronounced in reversible contexts. In nonreversible situations, the presence of SOV and VO orders becomes more significant.

The analysis of 93 locative sentences (see Table 2) revealed the following word order patterns:

Table 2. Sentence Order Distribution in Locative Constructions.

Topic		Count Sentence	%
Locative	OV	2	2.15
	SOV	25	26.88
	SVO	41	44.09
	OSV	25	26.88
	Total	93	100

The object-verb (OV) order appears in 2 sentences, making up 2.15% of the total. The subject-object-verb (SOV) order is observed in 25 sentences, accounting for 26.88%. The most common structure is subject-verb-object (SVO), found in 41 sentences, representing 44.09%. The object-subject-verb (OSV) order also occurs in 25 sentences, comprising 26.88%.

These findings suggest that while SVO is the most frequently used structure, the SOV and OSV orders are equally significant in locative sentences.

The analysis of 199 sentences (Table 3), categorized based on animacy, revealed the following patterns:

Table 3. Sentence Order Distribution Based on Animacy (Animate and Inanimate Objects).

Topic		Count Sentence	%
Animate	OV	56	53.33
	VO	49	46.67
	Total	105	100
Inanimate	OV	70	74.47
	VO	24	25.53
	Total	94	100

In sentences with animate objects, the preferred word order is object-verb (OV), found in 56 sentences (53.33%). However, the verb-object (VO) order is also common, appearing in 49 sentences (46.67%), indicating a relatively balanced distribution between the two structures.

In sentences with inanimate objects, the OV order is clearly dominant, occurring in 70 sentences (74.47%), while the VO order is significantly less frequent, found in only 24 sentences (25.53%).

These findings suggest that for inanimate objects, the OV order is strongly preferred, whereas for animate objects, both OV and VO orders are used with nearly equal frequency.

The analysis of 47 sentences (Table 4), categorized based on the heaviness of the object, revealed the following patterns:

Table 4. Sentence Order Distribution Based on Object Heaviness (Heavy and Nonheavy Objects).

Topic		Count Sentence	%
Heavy	OV	30	76.92
	VO	9	23.08
	Total	39	100
Nonheavy	OV	7	87.5
	VO	1	12.5
	Total	8	100

For heavy objects, the object-verb (OV) order is the preferred structure, occurring in 30 sentences (76.92%). The verb-object (VO) order is less frequent, appearing in only 9 sentences (23.08%).

In sentences with non-heavy objects, the OV order remains dominant, observed in 7 sentences (87.5%), while the VO order is recorded in just 1 sentence (12.5%).

These findings indicate that the OV order is consistently preferred, regardless of the object's heaviness, with an even stronger preference for non-heavy objects.

Following these observed patterns and structural rules, we developed a parallel corpus that aligns Kazakh text with its corresponding representation in KSL. This corpus was designed to accurately reflect the syntactic and grammatical variations identified in our analysis, ensuring a reliable foundation for further research in sign language processing and machine translation.

3.4. Algorithm and Pre-Processing

During the processing of input data, we transform the sentence in the following way:

$$F(S) = T_{temp}(R(S))$$

where:

- $R(S)$ performs word rearrangement.
- T_{temp} adds temporal markers.
  (1) Initial Sentence (S):
  $S = (w_1, w_2, \ldots, w_n)$—represents a sequence of words forming a sentence, where:
- w_i is the i-th word in the sentence.
  (2) Syntactic Dependency $f(w_i)$:
  Each word w_i associated with a syntactic role:
  $f(w_i) \in$ {subject, object, verb, adverbial}.
  (3) Classification of the word's animacy
  $a(w) \in$ {animate, inanimate}.
  (4) Word Rearrangement Function $R(S)$:
  $R(S)$ transforms the sentence structure based on specific conditions:
- If both the subject and the object are animate, they switch places with the verb.
- If the sentence includes a locative, the locative and the subject are swapped.
- If one participant (subject or object) is animate, the object switches places with the verb.
  (5) Tense Adjustment Function $T_{temp}(R(S))$:
  Temporal markers are added:
- For past tense: "past marker".

- For present tense: "present marker".
- For future tense: "future marker future marker".

During the preprocessing stage, each sentence is processed through a parser, which is responsible for analyzing sentence structure [22] to identify grammatical components and the relationships between words.

The algorithm for constructing the text processing framework (see Figure 2) consists of several steps.

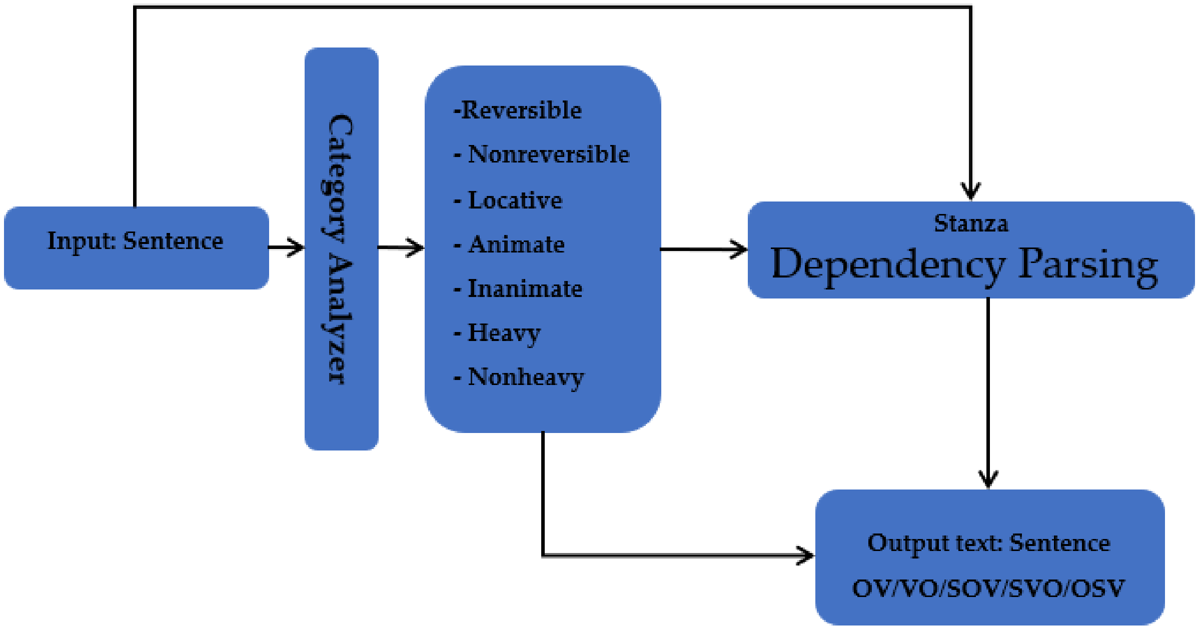


Figure 2. KSL translation algorithm.

At the first stage, input data are identified, including the type of text and the specific characteristics of interest.

At the second stage, data processing tools such as the Category Analyzer and Stanza are selected. The Category Analyzer is used to extract text categories and features, while Stanza performs syntactic analysis, identifying dependencies between words.

Following the analysis, each sentence is categorized based on specific criteria:

- A sentence is classified as locative if it contains "obl: loc". In this case, the sentence structure is converted to OSV.
- A sentence is considered reversible if it contains both a subject and an object. If both are identified as animate, the structure is converted to SVO.
- Animate classification follows similar rules as reversible sentences, but instead of checking for two animate elements, it only verifies the presence of one. If an animate element is found, the verb and object are swapped, forming an SVO structure.

At the next stage, the processed output data are generated based on the results from the previous steps.

3.5. Training

In this section, we focus on training our parallel text using the Transformer architecture and evaluating its performance in comparison with existing studies. The Transformer model, originally introduced in the paper *Attention is All You Need* [25], has proven to be one of the most effective models for machine translation.

Figure 3 presents the architecture of the Transformer, which is composed of two primary components: the encoder and the decoder. The encoder processes the input sequence of tokens, transforming them into context-sensitive vector representations through the application of multi-head attention, layer normalization, and feed-forward layers. The decoder then generates the output sequence by employing masked multi-head attention to manage previous tokens, while a separate multi-head attention mechanism incorporates information from the encoder. To compensate for the lack of recurrence in the architecture, positional encoding is utilized. The final output from the decoder is mapped to a probability distribution through a linear layer followed by a softmax function.

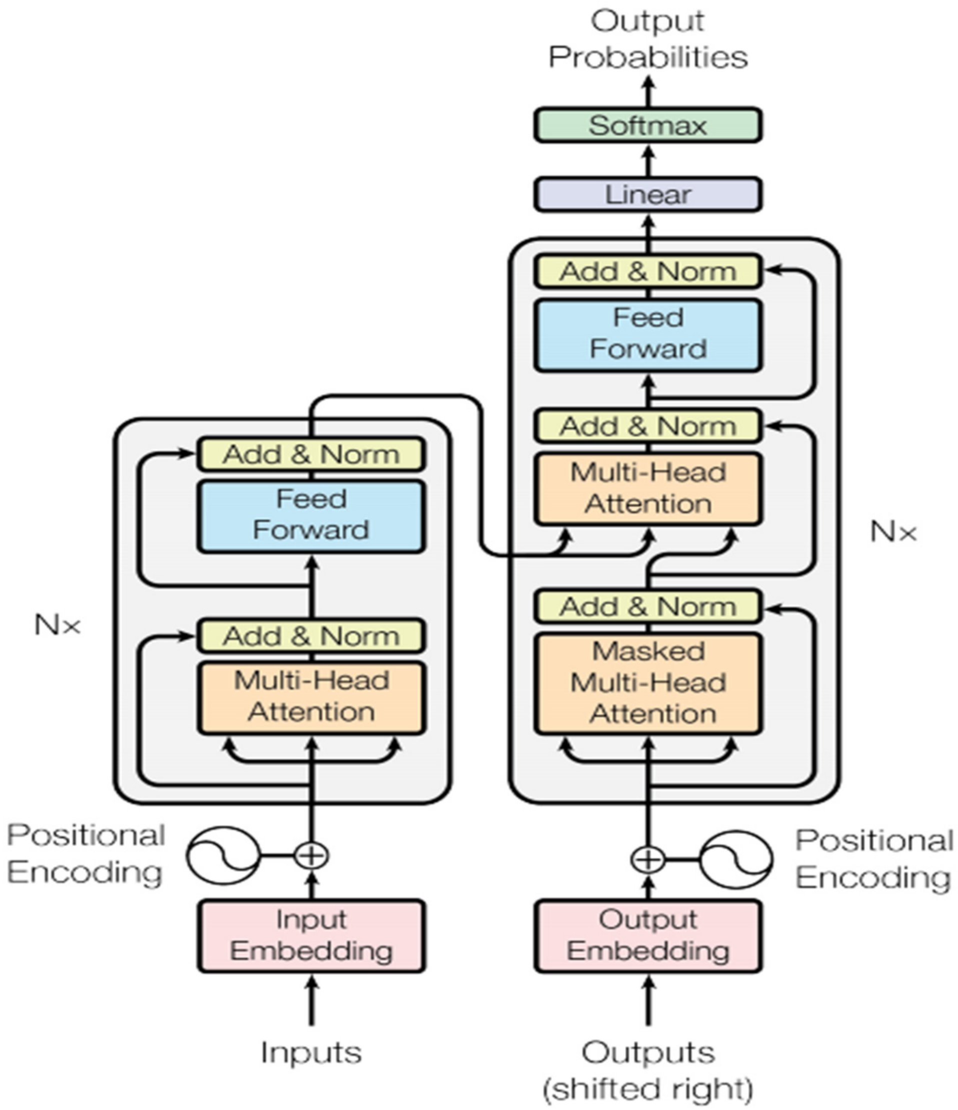


Figure 3. The Transformer—model architecture [26].

The Transformer model was chosen for this study due to its ability to efficiently process long sequences, its capacity for parallelizing data processing to accelerate learning, and its use of the attention mechanism to establish dependencies between distant elements in a sequence. These features make it particularly suitable for tasks involving complex linguistic structures, such as sign language translation.

The Transformer model used in this study was implemented in PyTorch 2.1. The training process was optimized using the Adam optimizer with hyperparameters set to $\beta 1 = 0.9$, $\beta 2 = 0.98$, and $= 10^{-9}$, following the configuration outlined in [27]. The model was trained with the following parameters:

- batch size = 64;

- dinner hid = 1024;
- dk = 64;
- dmodel = 512;
- dv = 64;
- dword vec = 512;
- dropout = 0.1;
- epochs = 50;
- maxtoken seq len = 59;
- nhead = 8;
- nlayers = 6;
- nwarmup steps = 4000;
- lrate = d − 0.5 modelmin (step − 0.5, step − 1.5 nwarmup steps).

3.6. Evaluation Metrics

We used BLEU (Bilingual Evaluation Understudy) to assess the quality of the generated text. BLEU is a widely used metric that evaluates machine-generated translations by comparing them to one or more human reference translations. The evaluation is based on the overlap of *n*-grams (sequences of n words) between the generated text and the reference translations, providing a quantitative measure of translation accuracy.

The BLEU score is calculated using the following Formulas (1)–(4):

$$Count_{clip}(n-gram) = \min\{Count(n-gram), Max\mathrm{Re}f Count(n-gram)\} \tag{1}$$

$$P_n = \frac{\sum_{C\in\{Candidates\}}\sum_{n-gram\in C} Count_{clip}(n-gram)}{\sum_{C'\in\{Candidates\}}\sum_{n-gram'\in C'} Count(n-gram')} \tag{2}$$

$$BP = \begin{cases} 1 & if\ c > r \\ e^{(1-r/c)} & if\ c < r \end{cases} \tag{3}$$

$$BLEU = BP \times \exp[(\sum_{n=1}^{N} w_n \log P_n)] \tag{4}$$

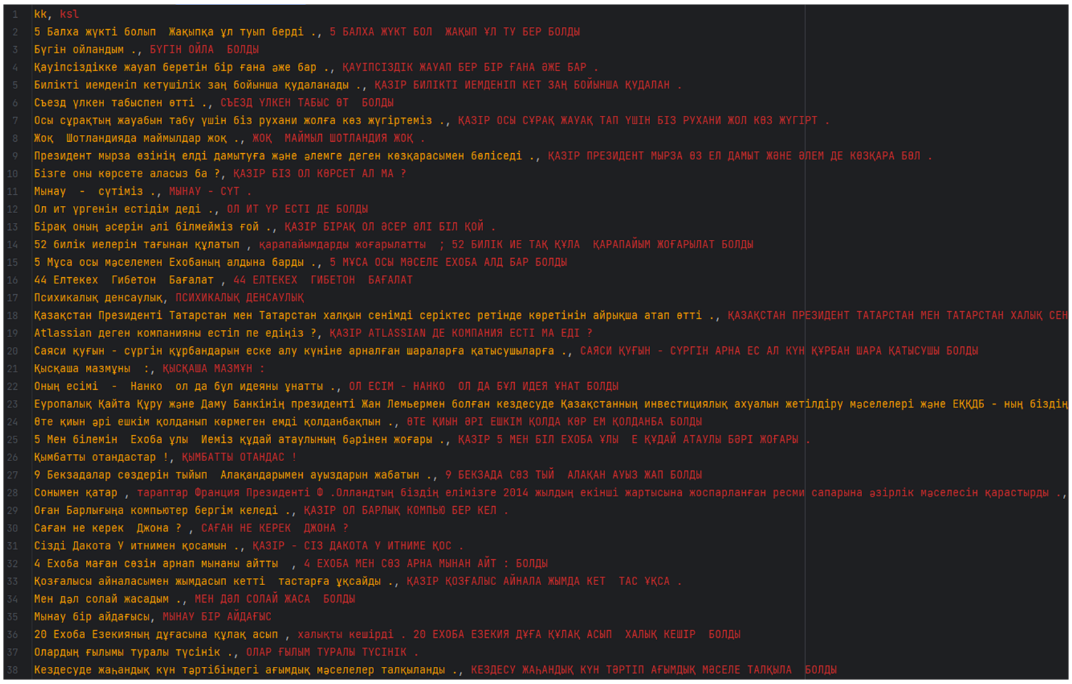
In this formula, the term "*n*-gram" refers to the frequency of *n*-grams appearing in the machine-generated translation that also occur in the reference translation. In this study, the value of n is set to 4, meaning that unigrams, bigrams, trigrams, and four-grams are taken into account. Additionally, the parameters used in the formula include (a) r, which represents the length of the reference translation; (b) c, which denotes the length of the machine-generated translation.

These elements contribute to an objective evaluation of translation quality by measuring the overlap between the generated and reference translations.

4. Results

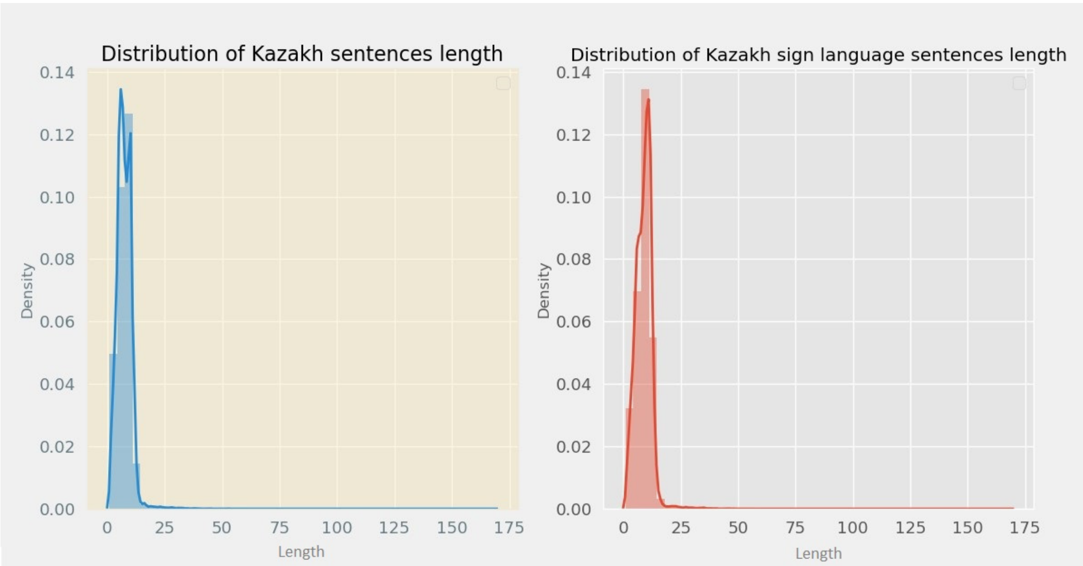
A study on the basic structure of KSL was conducted, providing insights into word order and the general structure of sign language. More than 10,000 sign language sentences were analyzed and classified into categories such as local case, animacy, reversibility, time, direction, and others.

The process of creating a parallel corpus consists of three key stages, each essential for optimizing translation and adaptation between Kazakh and KSL (see Figure 4).



```
kk, ksl
5 Балха жүкті болып  Жақыпқа ұл туып берді ., 5 БАЛХА ЖҮКТ БОЛ  ЖАҚЫП ҰЛ ТУ БЕР БОЛДЫ
Бүгін ойландым ., БҮГІН ОЙЛА  БОЛДЫ
Қауіпсіздікке жауап беретін бір ғана әже бар ., ҚАУІПСІЗДІК ЖАУАП БЕР БІР ҒАНА ӘЖЕ БАР .
Билікті иемденіп кетушілік заң бойынша қудаланады ., ҚАЗІР БИЛІКТІ ИЕМДЕНІП КЕТ ЗАҢ БОЙЫНША ҚУДАЛАН .
Съезд үлкен табыспен өтті ., СЪЕЗД ҮЛКЕН ТАБЫС ӨТ  БОЛДЫ
Осы сұрақтың жауабын табу үшін біз рухани жолға көз жүгіртеміз ., ҚАЗІР ОСЫ СҰРАҚ ЖАУАҚ ТАП ҮШІН БІЗ РУХАНИ ЖОЛ КӨЗ ЖҮГІРТ .
Жоқ  Шотландияда маймылдар жоқ ., ЖОҚ  МАЙМЫЛ ШОТЛАНДИЯ ЖОҚ .
Президент мырза өзінің елді дамытуға және әлемге деген көзқарасымен бөліседі ., ҚАЗІР ПРЕЗИДЕНТ МЫРЗА ӨЗ ЕЛ ДАМЫТ ЖӘНЕ ӘЛЕМ ДЕ КӨЗҚАРА БӨЛ .
Бізге оны көрсете аласыз ба ?, ҚАЗІР БІЗ ОЛ КӨРСЕТ АЛ МА ?
Мынау  -  сүтіміз ., МЫНАУ - СҮТ .
Ол ит үргенін естідім деді ., ОЛ ИТ ҮР ЕСТІ ДЕ БОЛДЫ
Бірақ оның әсерін әлі білмейміз ғой ., ҚАЗІР БІРАҚ ОЛ ӘСЕР ӘЛІ БІЛ ҚОЙ .
52 билік иелерін тағынан құлатып , қарапайымдарды жоғарылатты  ; 52 БИЛІК ИЕ ТАҚ ҚҰЛА  ҚАРАПАЙЫМ ЖОҒАРЫЛАТ БОЛДЫ
5 Мұса осы мәселемен Ехобаның алдына барды ., 5 МҰСА ОСЫ МӘСЕЛЕ ЕХОБА АЛД БАР БОЛДЫ
44 Елтекех  Гибетон  Бағалат , 44 ЕЛТЕКЕХ  ГИБЕТОН  БАҒАЛАТ
Психикалық денсаулық, ПСИХИКАЛЫҚ ДЕНСАУЛЫҚ
Қазақстан Президенті Татарстан мен Татарстан халқын сенімді серіктес ретінде көретінін айрықша атап өтті ., ҚАЗАҚСТАН ПРЕЗИДЕНТ ТАТАРСТАН МЕН ТАТАРСТАН ХАЛЫҚ СЕН
Atlassian деген компанияны естіп пе едіңіз ?, ҚАЗІР ATLASSIAN ДЕ КОМПАНИЯ ЕСТІ МА ЕДІ ?
Саяси қуғын - сүргін құрбандарын еске алу күніне арналған шараларға қатысушыларға ., САЯСИ ҚУҒЫН - СҮРГІН АРНА ЕС АЛ КҮН ҚҰРБАН ШАРА ҚАТЫСУШЫ БОЛДЫ
Қысқаша мазмұны  :, ҚЫСҚАША МАЗМҰН :
Оның есімі  -  Нанко  ол да бұл идеяны ұнатты ., ОЛ ЕСІМ - НАНКО  ОЛ ДА БҰЛ ИДЕЯ ҰНАТ БОЛДЫ
Еуропалық Қайта Құру және Даму Банкінің президенті Жан Лемьермен болған кездесуде Қазақстанның инвестициялық ахуалын жетілдіру мәселелері және ЕҚҚДБ - ның біздің
Өте қиын әрі ешкім қолданып көрмеген емді қолданбақпын ., ӨТЕ ҚИЫН ӘРІ ЕШКІМ ҚОЛДА КӨР ЕМ ҚОЛДАНБА БОЛДЫ
5 Мен білемін  Ехоба ұлы  Иеміз құдай атаулының бәрінен жоғары ., ҚАЗІР 5 МЕН БІЛ ЕХОБА ҰЛЫ  Е ҚҰДАЙ АТАУЛЫ БӘРІ ЖОҒАРЫ .
Қымбатты отандастар !, ҚЫМБАТТЫ ОТАНДАС !
9 Бекзадалар сөздерін тыйып  Алақандарымен ауыздарын жабатын ., 9 БЕКЗАДА СӨЗ ТЫЙ  АЛАҚАН АУЫЗ ЖАП БОЛДЫ
Сонымен қатар , тараптар Франция Президенті Ф .Олландтың біздің елімізге 2014 жылдың екінші жартысына жоспарланған ресми сапарына әзірлік мәселесін қарастырды .,
Оған Барлығына компьютер бергім келеді ., ҚАЗІР ОЛ БАРЛЫҚ КОМПЬЮ БЕР КЕЛ .
Саған не керек  Джона ? , САҒАН НЕ КЕРЕК  ДЖОНА ?
Сізді Дакота У итнимен қосамын ., ҚАЗІР - СІЗ ДАКОТА У ИТНИМЕ ҚОС .
4 Ехоба маған сөзін арнап мынаны айтты  , 4 ЕХОБА МЕН СӨЗ АРНА МЫНАН АЙТ : БОЛДЫ
Қозғалысы айналасымен жымдасып кетті  тастарға ұқсайды ., ҚАЗІР ҚОЗҒАЛЫС АЙНАЛА ЖЫМДА КЕТ  ТАС ҰҚСА .
Мен дәл солай жасадым ., МЕН ДӘЛ СОЛАЙ ЖАСА  БОЛДЫ
Мынау бір айдағысы, МЫНАУ БІР АЙДАҒЫС
20 Ехоба Езекияның дұғасына құлақ асып , халықты кешірді . 20 ЕХОБА ЕЗЕКИЯ ДҰҒА ҚҰЛАҚ АСЫП  ХАЛЫҚ КЕШІР  БОЛДЫ
Олардың ғылымы туралы түсінік ., ОЛАР ҒЫЛЫМ ТУРАЛЫ ТҮСІНІК .
Кездесуде жаһандық күн тәртібіндегі ағымдық мәселелер талқыланды ., КЕЗДЕСУ ЖАҺАНДЫҚ КҮН ТӘРТІП АҒЫМДЫҚ МӘСЕЛЕ ТАЛҚЫЛА  БОЛДЫ
```

Stage A—Translation of the Kazakh text to a gloss.



Stage B—Distribution of sentence length in KSL.

Characteristics	Corpus's Kazakh Set	Corpus's KSL Set
Sentences	84,707	84,707
Tokens	645,197	767,346
Max. sentence size	42 (words)	45 (words)
Min. sentence size	1 (words)	1 (words)
Vocabulary size	33,181	16,980

Stage C—Comparison of the characteristics of the Kazakh language and the KSL corpus.

Figure 4. Stages of processing text data (Stages A, B, C).

Stage A: This stage involves the translation of text into gloss, where corresponding phrases in Kazakh and KSL are aligned. This step serves as the foundation for further analysis and text alignment.

Stage B: Sentence length distribution in Kazakh and KSL is analyzed and visualized through graphs. The comparison reveals that both languages generally use short sentences; however, Kazakh exhibits a wider range of sentence lengths, whereas KSL follows a more compact distribution. As a result, longer sentences in Kazakh may need to be divided into shorter segments when translated into KSL.

Stage C: The characteristics of the Kazakh and KSL corpora are compared, including key metrics such as the number of sentences, number of tokens, maximum and minimum sentence lengths, and vocabulary size. These data provide valuable insights into the structural and complexity differences between the two languages, aiding in the accurate alignment and translation of parallel texts.

Together, these three stages form a structured approach to improving translation between Kazakh and KSL.

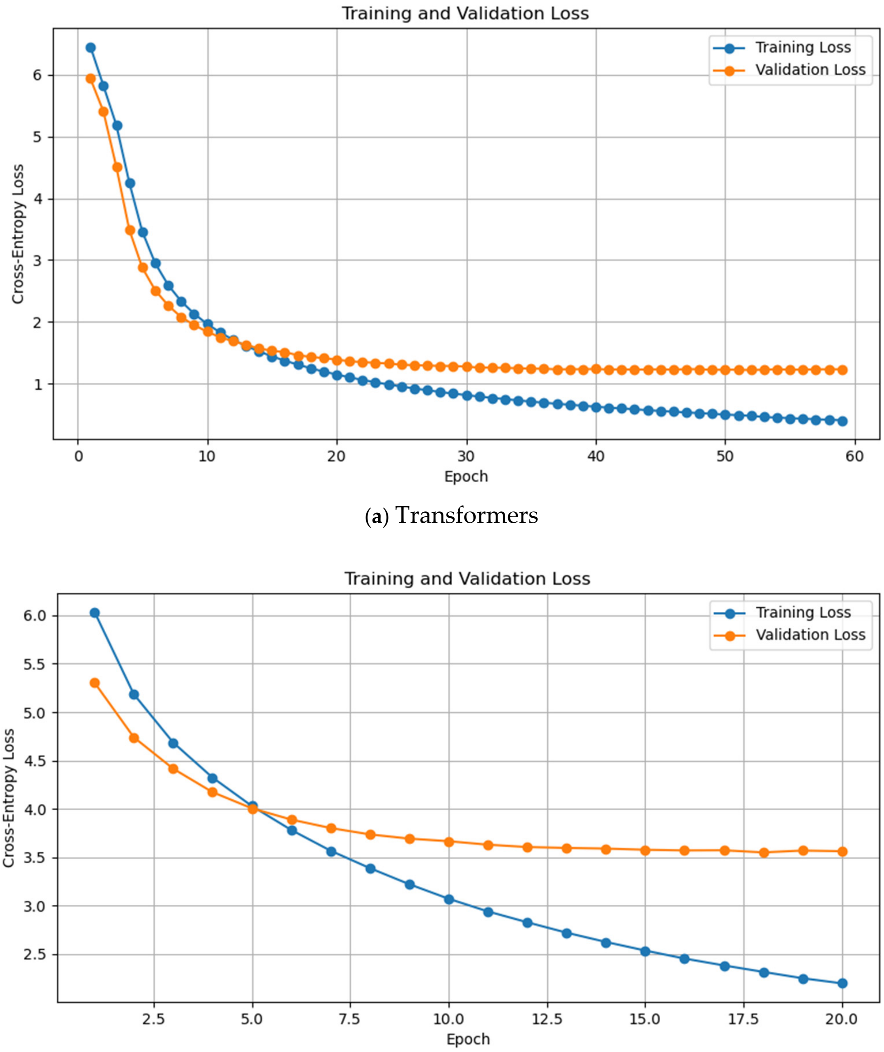
During this study, it became evident that the grammar of sign language differs significantly from that of spoken language, influencing the transformation rules for each of the designated categories. Based on these findings, a set of rules was developed, and a parser was created to convert Kazakh text into sign language text. Once the parser was completed, the model was trained on a dataset of over 100,000 sentences, sourced primarily from books and news articles. The training process was carried out using Transformer architecture.

In this study, the results of parallel corpus training were compared to assess the performance of the two most widely used machine translation system architectures (see Figure 5).

To assess the translation efficiency, the BLEU criterion was used. This algorithm evaluates the quality of machine-generated translations by comparing them to human translations. The quality assessment is based on the degree of similarity between the machine-generated text and the reference translation produced by a professional translator: "The more the machine translation matches the professional human translation, the higher its quality".—this is the key idea of BLEU. Figure 6 presents the BLEU scores for the Seq2seq and Transformer models. The Seq2seq model starts with an initial BLEU score of approximately 0.625, gradually improving and reaching a peak of around 0.675. In contrast, the Transformer model begins with a higher BLEU score of approximately 0.72. Consistent growth is observed throughout all training epochs, with the score reaching a maximum of approximately 0.775 in the fifth epoch. At every stage of training, the Transformer model outperforms the Seq2seq model, demonstrating superior translation accuracy.

A comparison of the various methods employed in the field of sign language translation reveals a spectrum of performance and limitations. The syntax-aware transformer-based approach for text-to-sign-language gloss achieves a BLEU score of 53.52%, a result that is enhanced by the incorporation of syntactic information but constrained by the scarcity of sign language data and the potential for performance degradation. The Progressive Transformers method for end-to-end sign language production demonstrates a marginal improvement with a BLEU score of 55.65%, yet it is hindered by high computational costs, risks of overfitting to specific datasets, challenges in generalising to diverse sign languages and dialects, and difficulties in aligning gestures with variations in textual input.

In contrast, the proposed method utilising a Transformer and Seq2seq model with a parallel corpus demonstrates significant improvement, achieving a BLEU score of 77.50%, which underscores its effectiveness, particularly for KSL (See Table 5). While this manuscript emphasises the innovation of incorporating specialised rules and a custom parser for KSL, further elaboration is required on the impact of these adaptations on the model's performance. Addressing these challenges necessitates an exploration of linguistic properties unique to KSL and the potential enhancement of the model with additional grammatical normalization strategies or hybrid approaches to improve fluency and coherence in translations.



(**a**) Transformers

(**b**) Seq2Seq

Figure 5. (**a**,**b**) Learning Process (Loss Function).

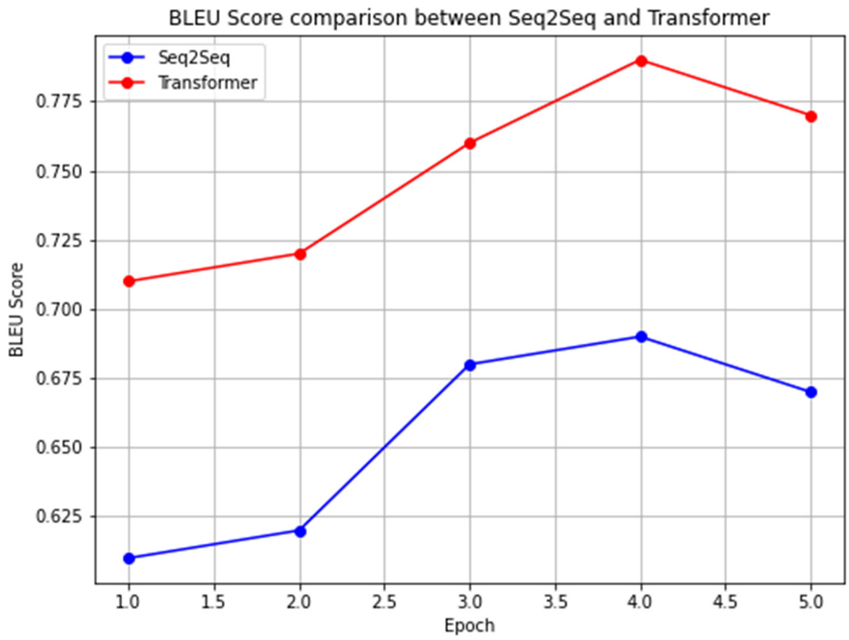


Figure 6. BLEU Score.

Table 5. Comparative table of methods and approaches.

Methods	Approaches	Metric (BLEU Score)	Limitations
Syntax-Aware Transformer-Based Machine Translation for Text-to-Sign-Language Gloss [28]	Incorporating Syntactic Information into Machine Translation Models	53.52%	The scarcity of SL data and potential performance degradation.
Progressive Transformers for End-to-End Sign Language Production Ben Saunders [29]	Progressive Transformers for End-to-End	55.65%	High computational costs, potential overfitting to specific datasets, challenges in generalizing to different sign languages and dialects, and difficulties in aligning gestures with variations in text input.
Our	Transformer and Seq2seq, parallel corpus	77.50%	Lack of uniform grammatical structures.

In the course of the present experiments, it was found that sentences comprising up to 42 words presented a considerable challenge to CNN models [9] due to their inability to capture long-range dependencies. Similarly, LSTMs [10], while demonstrating aptitude for capturing local patterns, were unable to grasp the global context of such sentences. By contrast, Transformers demonstrated superior performance, efficiently processing long sentences by means of their attention mechanism, which dynamically focuses on relevant parts of the input regardless of sentence length (See Table 6).

Table 6. Analysis of offer processing models.

Model Type	Strengths	Weaknesses	Performance on Long Sentences
CNN [9]	Effective at detecting local patterns	Challenges in understanding global context	Low
LSTM [10]	Good at sequential data, captures short-term dependencies	Struggles with long-range dependencies	Moderate
Transformer	Handles long sentences efficiently using attention mechanism	High computational cost	High

5. Discussion

This research provides valuable insights into the development of a machine translation system specifically designed for KSL, a linguistic field that has often been overlooked due to the limited availability of resources. By constructing a parallel corpus and analyzing the structural features of KSL, this study contributes to a deeper understanding of the complexities of sign language processing. It addresses both theoretical linguistic challenges and practical applications, particularly in assistive technologies aimed at supporting individuals with hearing and speech impairments in Kazakhstan.

A significant achievement of this study is the identification of core word order patterns in KSL. More than 500 natural language sentences were analyzed and categorized into seven groups: reversible, nonreversible, locative, animate, inanimate, heavy, and nonheavy, the research highlights key differences between KSL grammar and spoken Kazakh. These differences go beyond theoretical analysis, emphasizing the need to adapt existing natural language processing (NLP) frameworks to accommodate the structural demands of sign language. Unlike spoken languages that follow a linear syntax, sign languages like KSL operate within a visual–spatial modality, requiring more specialized and sophisticated modeling approaches.

A major technical breakthrough of this research is the development of a parser capable of converting Kazakh text into KSL glosses. This tool plays a crucial role in the creation of a comprehensive parallel corpus, which is essential for training machine learning models. This study further demonstrates the scalability and effectiveness of the system by applying Transformer-based machine translation models trained on diverse texts. The choice of Transformer models, known for their ability to capture complex contextual dependencies, aligns this research with recent advancements in NLP, leading to promising translation accuracy results as indicated by BLEU scores.

Despite these achievements, this study acknowledges several challenges. One of the most significant is the use of glosses as an intermediary step in the translation process. While glosses provide a practical method for building a parallel corpus, they often simplify the intricate visual–spatial grammar of sign language. To capture the full linguistic richness of KSL, future research could explore integrating video-based data and advanced multi-modal approaches. This would allow for a more accurate representation of non-manual markers such as facial expressions and body movements, which are essential for conveying meaning in KSL.

The evaluation metrics used in this study, particularly BLEU scores, offer a solid foundation for assessing translation quality. However, BLEU primarily focuses on n-gram overlap with reference translations, which may not fully capture the semantic depth and fluency needed for sign language translation. Future work could incorporate human evaluations or explore alternative metrics that prioritize meaning, context, and cultural relevance to provide a more comprehensive measure of translation quality.

Looking ahead, this research sets the stage for further advancements in sign language technology. Beyond translation, the methodologies and findings have the potential to impact real-time sign language interpretation, interactive learning tools, and inclusive communication systems. Expanding the parallel corpus to include conversational and domain-specific data would significantly enhance the system's adaptability and practical use.

The proposed approach to creating a parallel corpus for KSL offers a valuable framework that can be adapted in other countries, particularly within the post-Soviet region. This initiative brings several benefits, including language preservation, improved accessibility, and the promotion of intercultural communication.

One crucial aspect of this approach is its role in preserving endangered languages. Many regional sign languages face the risk of extinction, and the combination of corpus development and deep learning can serve as a method for documenting and revitalizing these languages, ensuring their survival for future generations.

Additionally, the creation of a parallel corpus facilitates intercultural communication both nationally and internationally. This is especially relevant in regions where sign languages share similarities or historical connections. The KSL project can serve as a model for collaborative efforts between countries to improve sign language research and accessibility.

A detailed analysis of KSL's grammatical structure also represents a significant contribution to sign language linguistics. Researchers working with other sign languages can adapt this methodology to their specific linguistic contexts, enabling more comparative studies and advancing the broader field of sign language research.

Technological advancements are another key aspect of this study. The use of deep learning, particularly the Transformer architecture, highlights the potential of advanced technologies in addressing challenges related to sign language processing. These innovations can be further adapted by other countries to suit their unique linguistic needs.

The methodologies developed for KSL can also be extended to other languages, particularly within the Commonwealth of Independent States (CIS) and former Soviet republics, where related sign languages exist or where research on sign languages remains underdeveloped. Collaborative projects and knowledge sharing among researchers from different countries could accelerate progress in this area.

This study addresses the resource gap for KSL, demonstrating the potential of NLP and machine learning in reducing communication barriers. Its findings contribute to advancements in sign language recognition while emphasizing the importance of inclusive, context-aware technological solutions. Ultimately, this research lays the groundwork for more accessible and inclusive systems, fostering greater communication opportunities for individuals with hearing and speech impairments.

6. Conclusions

This study addresses the issue of the lack of parallel data for KSL and presents methods for constructing a parallel corpus using glosses. This represents a significant step in the development of machine translation systems for individuals with hearing and speech impairments in Kazakhstan. Additionally, this study explores the structure of KSL by analyzing its word order in comparison to spoken languages. The data were examined to determine whether a basic word order can be identified in KSL.

At the final stage, the research focused on developing a parallel corpus to analyze KSL using NLP and machine learning techniques. A standard text processing pipeline was applied, incorporating tokenization, morphological analysis, and lemmatization, along with sentence categorization.

The parser was used to process over 100,000 sentences from books and news sources. A machine translation model was then built using the Transformer architecture, with training parameters outlined in the section "Parameters and Implementation Details". The results demonstrated high translation accuracy, highlighting the effectiveness of the proposed approach for processing KSL. This research contributes to the advancement of automated sign language processing systems, paving the way for new communication opportunities for individuals with hearing and speech impairments.

The development of a parallel corpus for KSL is a major achievement that expands research possibilities in the fields of machine translation and KSL processing, an area that has been largely unexplored in Kazakhstan. Existing machine learning methods were adapted to account for the unique grammatical and lexical characteristics of KSL, necessitating the creation of specialized rules, including a parser that converts Kazakh text into gloss format.

This study also includes an analysis of existing sign language translation systems, identifies their limitations, and demonstrates how the adapted approach enhances translation quality. The developed parallel corpus and accompanying tools, such as a mobile application for translating text into sign language, offer extensive opportunities for the deaf and hard-of-hearing community in Kazakhstan. These resources contribute to the development of educational materials, improved access to information, and greater social integration, emphasizing the importance of this work.

It is recommended that future research seek to evaluate the educational impact of the mobile application through the use of user surveys, the analysis of feedback on the App Store, and interviews with individuals with hearing and speech impairments. However, it is essential that these assessments are conducted only after full integration of the system into the mobile application. Such integration will facilitate the collection of meaningful user feedback, thereby ensuring that improvements are made based on actual experiences and needs.

Author Contributions: Conceptualization, A.Y. and M.S.; methodology, E.D.; software, B.S.; formal analysis, date base, M.O.; mathematical models, U.B.; resources, B.S.; algorithm, mathematic analysis, B.S. and U.B.; data curation, U.B.; writing—original draft preparation, M.S.; writing—review and editing, E.D. and U.B.; project administration, A.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR24992875).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article; further inquiries can be directed to the corresponding author.

Acknowledgments: We express our gratitude to Bayekeyeva Ainur (Translation Studies and Applied Linguistics Department, School of Liberal Arts, Maqsut Narikbayev University, 010000, Astana, Kazakhstan; abaekeeva@gmail.com) for her assistance in working on the article, specifically for the linguistic analysis, reviewing, and editing of the article's text.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ma, Y.; Zhou, G.; Wang, S.; Zhao, H.; Jung, W. SignFi: Sign Language Recognition Using WiFi. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–21. [CrossRef]
2. Bragg, D. Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective. In Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility, Pittsburgh, PA, USA, 28–30 October 2019; pp. 16–31. [CrossRef]
3. Rastgoo, R.; Kiani, K.; Escalera, S. Sign language recognition: A deep survey. *Expert Syst. Appl.* **2021**, *164*, 113794. [CrossRef]
4. Amrutha, B.L.; Harshitha, H.M.; Nisarga, G.P.; Sindhu, S.R.; Mrs, G.V. Sowmya. Speech to Sign Conversion Using Natural Language Processing. *Int. J. Sci. Res. Eng. Manag.* **2024**, *8*, 1–5. [CrossRef]
5. Sparsha, U.; Priyanka, M.; Mukthashree, S.; Kiran, K.N. System for Sign Language to Speech Conversion. In Proceedings of the 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bengaluru, India, 26–27 April 2024; pp. 1–4. [CrossRef]
6. De Coster, M.; Shterionov, D.; Van Herreweghe, M.; Dambre, J. Machine Translation from Signed to Spoken Languages: State of the Art and Challenges. *Univers. Access Inf. Soc.* **2024**, *23*, 1305–1331. [CrossRef]
7. Mukhanov, S.; Uskenbayeva, R.; Rakhim, A.A.; Akim, A.; Mamanova, S. Gesture Recognition of the Kazakh Alphabet Based on Machine and Deep Learning Models. *Procedia Comput. Sci.* **2024**, *241*, 458–463. [CrossRef]
8. Amirgaliyev, Y.; Ataniyazova, A.; Buribayev, Z.; Zhassuzak, M.; Urmashev, B.; Cherikbayeva, L. Application of Neural Networks Ensemble Method for the Kazakh Sign Language Recognition. *Bull. Electr. Eng. Inform.* **2024**, *13*, 3275–3287. [CrossRef]
9. Yerimbetova, A.; Sakenov, B.; Berzhanova, U.; Mukazhanov, N.; Daiyrbayeva, E.; Othman, M. Development of a Model of Kazakh Sign Language Recognition Based on Deep Learning Method. In Proceedings of the 2024 9th International Conference on Computer Science and Engineering (UBMK), Antalya, Turkey, 4–6 September 2024; pp. 822–827. [CrossRef]
10. Yerimbetova, A.; Kaidina, D.; Sakenov, B.; Daiyrbayeva, E.; Turdalyuly, M.; Berzhanova, U. Recognising Kazakh Sign Language with Mediapipe. In Proceedings of the 2024 9th International Conference on Computer Science and Engineering (UBMK), Antalya, Turkey, 4–6 September 2024; pp. 828–833. [CrossRef]
11. Othman, A.; Jemni, M. English-ASL Gloss Parallel Corpus 2012: ASLG-PC12. In Proceedings of the sign-lang@LREC 2012 Workshop, Istanbul, Turkey, 21–22 May 2012; European Language Resources Association (ELRA): Istanbul, Turkey, 2012; pp. 151–154.
12. Forster, J. Automatic Sign Language Recognition: From Video Corpora to Gloss Sentences. Master's Thesis, RWTH Aachen University, Aachen, Germany, 2020.
13. Hanke, T.; Schulder, M.; Konrad, R.; Jahn, E. Extending the Public DGS Corpus in Size and Depth. In Proceedings of the Sign-lang@LREC 2020 Workshop, Marseille, France, 11–16 May 2020; European Language Resources Association (ELRA): Marseille, France, 2020; pp. 75–82.
14. Kim, Y.; Kwak, M.; Lee, D.; Kim, Y.; Baek, H. Keypoint-based sign language translation without glosses. *arXiv* **2022**, arXiv:2204.10511. [CrossRef]
15. Mahmoud, H.E.; Elattar, M.A.; Darweesh, M.S. Joint Sign Language Recognition and Translation Using Keypoint Estimation. In Proceedings of the 2023 11th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC), Alexandria, Egypt, 18–20 December 2023; pp. 184–189.
16. Kim, Y.; Baek, H. Preprocessing for keypoint-based sign language translation without glosses. *Sensors* **2023**, *23*, 3231. [CrossRef] [PubMed]
17. Margalitadze, T.; Meladze, G.; Pourtskhvanidze, Z. English–Georgian parallel corpus and its application in Georgian lexicography. *Lexikos* **2022**, *32*, 158–176. [CrossRef]
18. Jiang, Z.; Göhring, A.; Moryossef, A.; Sennrich, R.; Ebling, S. SwissSLi: The Multi-parallel Sign Language Corpus for Switzerland. In Proceedings of the 2024 Joint International Conference on Computational Linguistics and Language Resources and Evaluation (LREC-COLING 2024), Torino, Italy, 20–25 May 2024; pp. 15448–15456.
19. Stokoe, W.C. Sign Language Structure. *Annu. Rev. Anthropol.* **1980**, *9*, 365–390. [CrossRef]

20. Prillwitz, S.; Zienert, H. Hamburg Notation System for Sign Language Research: Development of a Sign Writing with Computer Application. In *Current Trends in European Sign Language Research*; Prillwitz, S., Vollhaber, T., Eds.; Signum: Hamburg, Germany, 1990; pp. 355–379.
21. Sutton, V.; Gleaves, R. *SignWriter—The World's First Sign Language Processor*; Deaf Action Committee for SignWriting: La Jolla, CA, USA, 1995.
22. Kimmelman, V. Word order in Russian sign language. *Sign Lang. Stud.* **2012**, *12*, 414–445. [CrossRef]
23. Zaitseva, G.L. *Gesture and Word: Scientific and Methodological Articles*; Moscow Press: Moscow, Russia, 2006; 631p.
24. Qi, P.; Zhang, Y.; Zhang, Y.; Bolton, J.; Manning, C.D. Stanza: A Python natural language processing toolkit for many human languages. *arXiv* **2020**, arXiv:2003.07082. [CrossRef]
25. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
26. Mars, M. From word embeddings to pre-trained language models: A state-of-the-art walkthrough. *Appl. Sci.* **2022**, *12*, 8805. [CrossRef]
27. Velay, M.; Daniel, F. Seq2seq and multi-task learning for joint intent and content extraction for domain-specific interpreters. *arXiv* **2018**, arXiv:1808.00423. [CrossRef]
28. Gómez, S.E.; Fabra, S.U.P.; McGill, E.; Saggion, H. Syntax-Aware Transformers for Neural Machine Translation: The Case of Text to Sign Gloss Translation. In Proceedings of the 14th Workshop on Building and Using Comparable Corpora (BUCC 2021), Online, 18 June 2021; pp. 18–27. [CrossRef]
29. Saunders, B.; Camgoz, N.C.; Bowden, R. Progressive transformers for end-to-end sign language production. In *Computer Vision–ECCV 2020 Proceedings of the 16th European Conference, Glasgow, UK, 23–28 August 2020*; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 687–705. [CrossRef]