# BETO-RRF: A Hybrid Template- and Rule-Based Framework for Mexican Sign Language Gloss Translation: A Comparison with Seq2Seq Models

Marcelo Alejandro Huerta-Espinoza[1,*,†], Metztli Ramírez-González[2,†]

[1]Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Santa María Tonantzintla, Puebla , Mexico
[2]Posgrado en Ciencia e Ingeniería de la Computación, Universidad Nacional Autónoma de México (UNAM), CDMX , México

### Abstract

In the context of Mexican Sign Language (MSL), glosses constitute an important intermediate representation for the computational modeling and understanding of sign language. Consequently, gloss-based representations have gained increasing relevance within the Natural Language Processing (NLP) field, particularly for developing systems capable of accurately translating between Spanish and gloss sequences. In this paper, we describe our participation in the MSLG-SPA 2026 shared task and explore different strategies for automatic bidirectional translation between Mexican Sign Language glosses and Spanish. We evaluate three translation configurations: two transformer-based Seq2Seq approaches and a hybrid template- and rule-based framework specifically designed for gloss translation. The hybrid framework combines contextual text representations, template prediction, and linguistic rules to generate the translated output. Through these experiments, we analyze the behavior of general neural translation models and a linguistically guided approach for gloss-based translation. The model results demonstrate that generating natural Spanish is a more challenging task, where Seq2Seq approaches achieve superior performance. In contrast, hybrid template-guided methods appear more effective for generating gloss representations, likely due to the simplified and structured syntactic patterns commonly observed in gloss sequences.

### Keywords

MSL, Glosses, Translation, T5, BETO, Random Forest, Translation Rules

## 1. Introduction

Oral communication is one of the most common mechanisms through which people participate in society. However, for individuals who are born with or acquire hearing impairments, communication barriers can significantly reduce opportunities for social participation, education, and professional development. To address this need, Deaf communities have developed their own linguistic systems, known as sign languages [1, 2, 3].

Mexican Sign Language (MSL) is the natural language used by the Deaf community in Mexico. Unlike spoken languages, MSL is expressed through a visual–gestural modality and possesses its own grammatical structure and linguistic organization. It combines manual components (hand configurations, movements, and spatial organization) with non-manual features such as facial expressions and body posture, all of which contribute to meaning. In addition, manual alphabets are used to represent proper nouns, locations, or concepts that do not yet have an established sign [1, 3].

Since sign languages do not possess a standardized written form equivalent to spoken languages, linguistic studies commonly rely on glosses as an intermediate textual representation. Glosses are conventionally written in uppercase letters and provide a symbolic description of signs and their linguistic structure. Their use facilitates the analysis of sign language independently from spoken language structures and enables the study of syntactic, semantic, and morphological phenomena without assuming a direct one-to-one correspondence between signs and words[3].

*IberLEF 2026, September 2026, León, Spain*
*Corresponding author.
†These authors contributed equally to this work.
marcelo.huerta@inaoep.mx (M. A. Huerta-Espinoza); metztli.ramirez@unam.edu (M. Ramírez-González)
0009-0002-5040-9640 (M. A. Huerta-Espinoza); 0009-0007-4330-4476 (M. Ramírez-González)


CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2026-09-29

Beyond their use in linguistic studies, glosses have begun to gain relevance as a computational representation for the development of language technologies aimed at sign languages. Although educational resources and tools have been developed to support the learning and dissemination of MSL, including dictionaries [2], digital glossaries[4], and some initial proposals for translation software[5], the incorporation of sign languages into modern automatic translation systems remains a relatively recent and still underexplored research area. In this context, the development of models capable of translating between glosses and natural language represents a step toward more inclusive and accessible technologies.

To promote research in this direction, the MSLG-SPA 2026 shared task, organized within the framework of IberLEF [6], focuses on automatic translation between Mexican Sign Language Glosses (MSL-G) and Spanish in a bidirectional setting [7]. Unlike conventional machine translation, this task presents additional challenges due to the visual–gestural nature of sign languages, the absence of standardized writing systems, and structural differences between gloss sequences and natural language.

The shared task is divided into two complementary subtasks:

1. *MSL Gloss to Spanish (MSLG2SPA)*, where systems generate fluent and semantically accurate Spanish sentences from gloss sequences, and
2. *Spanish to MSL Gloss (SPA2MSLG)*, where systems produce gloss sequences that preserve the semantic content and grammatical structure of MSL.

In this paper, we describe our participation in the MSLG-SPA 2026 shared task. We explore three translation configurations: (1) a conventional multilingual transformer model, (2) a transformer model adapted using MSL gloss data, and (3) a hybrid template- and rule-based framework designed specifically for gloss translation. The proposed framework consists of four sequential stages: (i) contextual feature extraction using BETO, (ii) template prediction through a Random Forest classifier, (iii) application of template-specific linguistic rules, and (iv) generation of the translated sentence.

The remainder of this paper is organized as follows. Section 2 presents previous efforts related to gloss-based translation approaches. Section 3 describes the experimental setup and development results. Section 4 reports the official results obtained in the MSLG-SPA 2026 shared task. Finally, Section 5 concludes the paper.

## 2. Related work

In Mexico, hearing impairment affects a significant portion of the population and continues to pose challenges for social participation and access to communication. Since MSL is considered a minority language in a predominantly Spanish-speaking environment, communication barriers may restrict educational, professional, and social integration. This scenario has motivated the development of technological solutions aimed at reducing accessibility gaps and facilitating communication [8, 9].

Early technological efforts for Mexican Sign Language have primarily focused on visual sign recognition systems, where the objective is to capture and interpret hand movements directly from visual input. One representative approach proposed a bidirectional translation system for MSL in healthcare environments, integrating sign recognition and generation to support communication in primary care services. Their system employed a Microsoft Kinect sensor to capture trajectories and image information and used Hidden Markov Models (vc2) for real-time recognition, reporting promising recognition performance[10]. Another line of work explored the development of software for desktop and mobile devices capable of interpreting MSL while incorporating object detection mechanisms to enrich vocabulary acquisition and facilitate communication between Deaf and hearing users[5].

Although these approaches have demonstrated the feasibility of visual sign translation, recent work has increasingly explored gloss-based representations as an intermediate layer between sign language and natural language processing techniques. Glosses allow sign language structures to be represented symbolically, enabling the application of translation methods originally designed for text processing.

Among gloss-based approaches, previous work proposed an architecture for Spanish-to-MSL gloss translation combining two paradigms: a rule-based system using automata and a neural approach based on deep learning. This work highlighted the importance of considering the grammatical organization of MSL during translation, discussing constituent order and linguistic variability in gloss generation[9]. More recent efforts evaluated transformer-based neural machine translation models using pretrained architectures and gloss corpora to study the applicability of modern sequence-to-sequence methods for MSL translation[11].

These latter approaches particularly influenced the design of our work. On one hand, transformer-based translation models demonstrate the ability to learn translation mappings directly from data. On the other hand, rule-based approaches offer an opportunity to incorporate explicit linguistic structure into the translation process while maintaining lower computational complexity compared to transformer-based methods. Additionally, their explicit modeling of linguistic patterns may facilitate the representation of the simplified and structured syntax typically found in gloss sequences. Motivated by these observations, we propose exploring a hybrid framework that combines learned representations with automatically induced translation rules. Rather than relying exclusively on static linguistic templates or purely data-driven learning, our objective is to construct a more structured translation process capable of capturing recurrent gloss patterns while maintaining interpretability. Such explicit representations may also contribute to the development of introductory educational resources and facilitate understanding of gloss organization for new learners of MSL.

## 3. Proposed Models and Experiments

The MSLG-SPA task consists of bidirectional translation between MSL glosses and Spanish through two complementary subtasks: Gloss-to-Spanish and Spanish-to-Gloss. For each subtask three models were designed and trained, in this section we explain the three architectures employed for each task. Every architecture was used for both subtasks.

### 3.1. Dataset

The experiments were conducted using a parallel MSLG–Spanish dataset composed of 490 paired instances. Each sample contains an identifier, a gloss sequence, and its corresponding Spanish translation. The dataset includes diverse linguistic structures such as declarative sentences, interrogative forms, temporal expressions, and negations, providing variability in sentence organization and communicative intention. The gloss sequences contain an average length of 4.7 tokens, whereas Spanish sentences present an average of 6.4 words, reflecting the more compact nature of gloss representations compared to natural language expressions. The dataset was employed for both Spanish-to-Gloss and Gloss-to-Spanish translation tasks, enabling the evaluation of bidirectional translation approaches under a low-resource setting.

Prior to training, samples with missing or empty values were removed, and text normalization procedures were applied to reduce inconsistencies in punctuation and formatting. Subsequently, the dataset was divided into training and validation subsets using a 90%/10% split with a fixed random seed (42) to ensure reproducibility.

### 3.2. T5-based models

The first two models were developed using transformer-based Seq2Seq architectures for translation tasks. The first model employed was T5 Base Spanish [12], a Spanish-adapted version of the Text-to-Text Transfer Transformer (T5) architecture. T5 is an encoder-decoder transformer model designed under a unified text-to-text framework, where all natural language processing tasks are reformulated as text generation problems. This design enables the model to transfer contextual knowledge learned during large-scale pretraining to downstream tasks through fine-tuning. Following the T5 paradigm, translation tasks were formulated as instruction-based text generation problems by prepending the input with task-specific prompts. For the MSL gloss to Spanish task, the prompt “traducir glosa a español:” was used, whereas for the Spanish to MSL gloss task, the prompt “traducir español a glosa:”

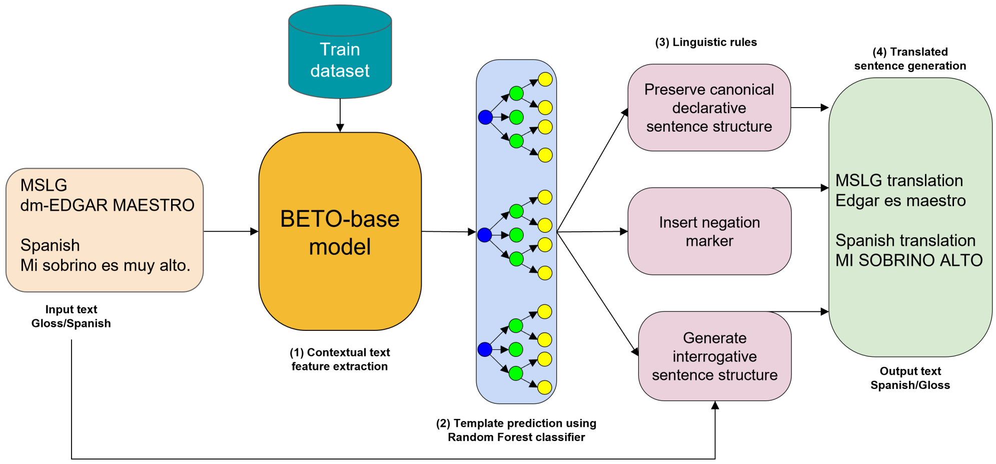


**Figure 1:** BETO-RRF architecture.

was employed. This prompt-based formulation guides the model toward the desired translation direction while maintaining a unified text-to-text learning framework. The model was subsequently fine-tuned using the complete training dataset, allowing all parameters to adapt to each translation task.

The second model utilized was Esp to LSM BARTO [13], a transformer model previously fine-tuned on a Spanish–Mexican Sign Language gloss corpus and built upon the Spanish BART architecture. Similar to the first approach, this model was further fine-tuned using the current dataset to adapt its learned representations to the target domain. The same prompt-based strategy was employed across both translation directions to maintain consistency between architectures and ensure comparable experimental settings. Since the model had already been exposed to Spanish–MSL translation patterns, it incorporates prior knowledge specialized in sign language translation tasks.

Both models were trained under a supervised learning setting using the same training data and optimization procedure to ensure a fair comparison. Fine-tuning was performed for 50 epochs using the AdamW optimizer with a learning rate of $1 \times 10^{-5}$, a batch size of 2 for both training and validation, and a weight decay of 0.01. Input and target sequences were truncated to a maximum length of 64 tokens. During validation and inference, translations were generated using beam search with four beams and a maximum output length of 64 tokens. Model selection was based on the highest validation chrF score, and the corresponding checkpoint was retained for the final evaluation. The objective of employing these architectures was to compare the behavior of a general Spanish pretrained Seq2Seq model against a model with previous specialization in Spanish-to-MSL translation, exploring the effect of prior domain-specific knowledge on translation performance.

### 3.3. BETO-RRF model

The third proposed architecture was the BETO-Rule Random Forest model (BETO-RRF), a hybrid template- and rule-based framework designed for gloss translation. Figure 1 illustrates the architecture of the BETO-RRF model. The framework is composed of four sequential stages: (1) contextual text feature extraction using BETO [14], (2) template prediction through a Random Forest classifier, (3) application of template-specific linguistic rules, and (4) generation of the translated sentence.

In the first stage, contextual embeddings are extracted from gloss sequences using the BETO model. BETO was fine-tuned using the complete training dataset to adapt its representations to the linguistic characteristics of the translation task. The extracted embeddings provide semantic and contextual information about the input sequence and are later used for template classification.

The objective of the second stage is to infer a linguistic template representing the underlying syntactic structure or communicative intention of the sentence. Rather than directly generating a translated sentence, the proposed framework decomposes the translation process into structural prediction followed by linguistic realization. This decomposition reduces the complexity of direct gloss-to-text mapping and facilitates the application of rule-based transformations for sentence generation. Table 1 summarize the templates used for the Spanish-to-gloss and gloss-to-Spanish translation tasks, respectively.

**Table 1**
Template categories used for sentence structure prediction. Each template represents a high-level linguistic pattern inferred from target sentences.

| Template | Description | Example Sentence | Function |
|---|---|---|---|
| TPL_DECL_SIMPLE | Simple declarative sentence | “Yo como pan” | Statement |
| TPL_NEGACION | Negative sentence structure | “Yo no quiero ir” | Negation |
| TPL_DECL_PASADO | Past tense declarative sentence | “Yo fui ayer” | Temporal (Past) |
| TPL_DECL_FUTURO | Future tense declarative sentence | “Yo viajaré mañana” | Temporal (Future) |
| TPL_WH_DONDE | Location-based interrogative sentence | “¿Dónde está María?” | WH-question |
| TPL_WH_QUIEN | Person-based interrogative sentence | “¿Quién vino?” | WH-question |
| TPL_WH_QUE | Object/event interrogative sentence | “¿Qué ocurrió?” | WH-question |
| TPL_WH_CUANDO | Time-based interrogative sentence | “¿Cuándo llegaste?” | WH-question |

Using the contextual features extracted by BETO, a Random Forest classifier was trained to predict the most suitable template for each input sequence. Random Forest was selected due to its robustness under limited-data scenarios, where training deep neural architectures may be constrained by the reduced availability of annotated examples. Furthermore, Random Forest models provide efficient non-linear decision boundaries while maintaining relatively low computational requirements.

After template prediction, a linguistic realization stage is performed through template-specific rules. These rules combine lexical mappings between gloss tokens and Spanish words with syntactic transformations associated with each template, Table 2 shows some examples of the transformations rules. Depending on the predicted structure, the generation process introduces linguistic modifications such as negation, interrogative forms, or temporal markers. Consequently, sentence generation becomes an interpretable process in which outputs can be traced to explicit structural transformations rather than solely relying on latent neural representations.

**Table 2**
Examples of template-guided linguistic realization rules applied during Gloss-to-Spanish sentence generation.

| Template(s) | Applied Rule | Generated Example |
|---|---|---|
| TPL_DECL_SIMPLE | Preserve canonical declarative structure after lexical mapping | YO COMER PAN → “Yo como pan” |
| TPL_NEGACION | Insert negation markers and reorganize lexical elements | YO NO IR → “Yo no voy” |
| TPL_DECL_PASADO,<br>TPL_DECL_FUTURO | Introduce temporal markers associated with past or future events | YO AYER ESCUELA IR → “Yo fui a la escuela ayer”<br>YO MANANA VIAJAR → “Yo viajaré mañana” |
| TPL_WH_DONDE,<br>TPL_WH_QUIEN,<br>TPL_WH_QUE,<br>TPL_WH_CUANDO | Generate interrogative structures according to semantic intention (location, person, object, or time) | MARIA ESTAR → “¿Dónde está María?”<br>LLEGAR → “¿Quién llegó?”<br>PASAR → “¿Qué pasó?”<br>LLEGAR → “¿Cuándo llegó?” |

After the template prediction stage, the final translated sentence is generated through a template-guided rule-based realization process. Initially, gloss tokens are transformed into their corresponding Spanish lexical representations using a predefined gloss-to-word mapping. However, this direct lexical substitution is insufficient to produce grammatically coherent sentences due to differences in structure between gloss representations and natural Spanish.

To address this limitation, the proposed framework applies a set of linguistic rules associated with the predicted template. These rules reorganize lexical elements according to the inferred syntactic structure and introduce specific transformations when required. For example, declarative templates preserve a canonical sentence structure, while negative templates incorporate negation markers. Similarly, interrogative templates modify sentence organization according to their semantic intention (e.g., person, location, object, or time questions), and temporal templates introduce markers associated with past or future events.

Consequently, sentence generation is decomposed into two complementary processes: lexical mapping and structural realization. This decomposition reduces the complexity of direct gloss-to-text translation while improving interpretability, since generated outputs can be traced back to explicit linguistic transformations guided by the predicted template rather than relying exclusively on latent neural representations. Such an approach may be particularly advantageous in low-resource scenarios, where limited training data constrains the performance of fully generative translation architectures.

### 3.4. Evaluation Metrics

To evaluate translation quality, multiple complementary metrics were employed, including BLEU, METEOR, chrF, and COMET. BLEU (Bilingual Evaluation Understudy) measures n-gram overlap between generated and reference sentences, providing an estimate of lexical similarity between translations. Higher BLEU scores indicate greater correspondence in token sequences; however, the metric may be sensitive to exact word matching and less effective in capturing semantic equivalence.

METEOR was additionally used to assess translation quality by considering lexical variations through stemming and synonym matching, allowing a more flexible comparison between generated and reference outputs. Consequently, METEOR may better reflect semantic similarity than strict n-gram matching approaches.

The chrF metric evaluates translation quality at the character level through F-score computation over character n-grams. This metric is particularly useful for morphologically rich languages or structured outputs such as gloss sequences, where preserving character patterns and token structure may be more informative than exact word-level overlap.

For the MSLG-to-Spanish task, COMET was also employed as a semantic evaluation metric. COMET is based on neural quality estimation models and measures the adequacy and fluency of generated translations by comparing source, hypothesis, and reference sentences. Unlike lexical overlap metrics, COMET aims to approximate human judgments of translation quality.

Since individual metrics capture different aspects of translation performance, an aggregated *Task Score* was computed to provide an overall comparison among models. The Task Score corresponds to the average of z-score normalized evaluation metrics, reducing scale differences between metrics and allowing a unified assessment of translation quality. Higher Task Score values indicate better overall performance across the evaluated criteria.

## 4. Results

The general results are presented in the task overview [7]; however, the performance obtained by our proposed models is summarized in this section. Table 3 presents the results for the Spanish-to-MSLG translation subtask. Overall, the BETO-RRF model achieved the best general performance according to the aggregated Task Score, outperforming the transformer-based approaches despite being a hybrid framework combining contextual embeddings, template classification, and rule-based generation.

More specifically, BETO-RRF obtained the highest chrF score (53.25), suggesting a better preservation of character-level patterns and structural similarity between generated and reference gloss sequences. This behavior may be associated with the template-guided generation process, which explicitly models sentence structures commonly observed in gloss representations. In contrast, the Esp-to-LSM BARTO model achieved the highest BLEU (11.99) and METEOR (0.3451) scores, indicating a stronger ability to reproduce token-level overlaps with reference sequences. Nevertheless, its lower chrF performance suggests reduced consistency in preserving fine-grained structural characteristics.

An additional factor that may explain the competitive performance of BETO-RRF is the architectural alignment between the model design and the target task. While T5-base Spanish and Esp-to-LSM BARTO are transformer architectures primarily optimized for natural language generation, particularly Spanish text generation tasks, BETO-RRF explicitly incorporates linguistic templates and rule-based transformations tailored to gloss generation. Since gloss sequences tend to exhibit reduced grammatical complexity and more structured syntactic patterns compared to natural language sentences, the template-guided approach may better capture these characteristics under low-resource conditions.

Furthermore, the T5-base Spanish model obtained the lowest overall performance across most metrics. A possible explanation is the limited size of the available parallel corpus, which may constrain the effective adaptation of fully generative transformer architectures. Under low-resource settings, approaches incorporating explicit linguistic priors, such as templates and rule-based transformations, may offer advantages by reducing the complexity of direct sequence generation.

These findings suggest that hybrid approaches integrating contextual embeddings with template-guided generation can provide competitive performance for sign language gloss translation, particularly when training data is scarce and the target representation follows structured linguistic patterns.

**Table 3**
Performance comparison for the Spanish-to-MSLG translation task. Task Score was computed as the average standardized performance across evaluation metrics. Best values are highlighted in bold.

| Model | Task Score | BLEU | METEOR | chrF |
|---|---|---|---|---|
| BETO-RRF | **-0.4964** | 5.1608 | 0.3447 | **53.2551** |
| Esp to LSM BARTO | -0.6149 | **11.9927** | **0.3451** | 36.9640 |
| T5 base Spanish | -1.7874 | 4.5939 | 0.1938 | 24.5633 |

Table 4 presents the results obtained for the MSLG-to-Spanish translation subtask. In contrast to the Spanish-to-MSLG scenario, the best overall performance was achieved by the T5-base Spanish model according to the aggregated Task Score, obtaining the highest BLEU (14.31), METEOR (0.3496), and COMET (0.6768) scores. These results suggest a greater capability to generate fluent and semantically coherent Spanish sentences, which is expected given that transformer-based sequence-to-sequence models are primarily optimized for natural language generation tasks.

**Table 4**
Performance comparison for the MSLG-to-Spanish translation task. Task Score was computed as the average standardized performance across evaluation metrics. Best values are highlighted in bold.

| Model | Task Score | BLEU | METEOR | chrF | COMET |
|---|---|---|---|---|---|
| BETO-RRF | -0.9943 | 2.1786 | 0.2208 | **44.6839** | 0.6490 |
| Esp to LSM BARTO | -1.5720 | 6.0448 | 0.2015 | 21.7692 | 0.4901 |
| T5 base Spanish | **-0.7019** | **14.3119** | **0.3496** | 36.7863 | **0.6768** |

The superior performance of T5-base Spanish in this subtask may be explained by the increased complexity of generating natural Spanish compared to producing gloss sequences. While gloss representations often follow simplified and structured patterns, Spanish generation requires handling richer grammatical structures, lexical variability, and contextual dependencies. Consequently, fully generative transformer architectures appear more suitable for producing fluent natural language outputs.

Despite not achieving the highest overall Task Score, the BETO-RRF model obtained the best chrF score (44.68), indicating a stronger preservation of character-level patterns and structural similarities between generated and reference sentences. This behavior suggests that template-guided approaches may still capture relevant structural information during translation, although their capacity to model linguistic variability remains limited compared to generative transformers.

Additionally, the lower performance observed for the Esp-to-LSM BARTO model may be associated with architectural misalignment with the evaluated task. Since the model was originally designed for sign language-related generation tasks, its adaptation to natural Spanish sentence generation may be constrained under the available low-resource conditions.

Overall, these findings indicate that template-guided hybrid approaches may be advantageous for generating structured outputs such as glosses, whereas transformer-based generative models exhibit greater effectiveness when the target modality corresponds to natural language sentences. This observation highlights the importance of aligning model architecture with the linguistic characteristics of the translation target.

## 5. Conclussion

Bidirectional translation between Mexican Sign Language glosses and Spanish represents an important yet challenging task due to the substantial linguistic differences between structured gloss representations and natural language sentences. While glosses typically follow simplified syntactic patterns and reduced grammatical variability, Spanish generation requires modeling richer linguistic structures, contextual dependencies, and greater lexical diversity. Consequently, translation performance may strongly depend on the alignment between model architecture and target representation.

The experimental results suggest that the proposed BETO-RRF framework constitutes a competitive alternative under low-resource conditions, particularly for the Spanish-to-MSLG subtask. The combination of contextual embeddings, template prediction, and rule-based linguistic realization appears effective for generating structured outputs such as gloss sequences, where explicit syntactic transformations can reduce the complexity of direct sequence generation. In this scenario, incorporating linguistic priors through templates may provide advantages over fully generative architectures.

However, for the MSLG-to-Spanish translation task, transformer-based models such as T5-base Spanish achieved superior overall performance. This behavior may be explained by the increased complexity associated with natural language generation, where fluent sentence production requires modeling semantic relationships, contextual nuances, and grammatical variability beyond predefined structural templates. Therefore, while template-guided hybrid approaches may be advantageous for generating structured representations, generative transformer architectures remain more suitable for producing coherent natural language outputs.

Overall, these findings highlight that no single architecture is universally optimal for bidirectional sign language translation. Instead, model effectiveness depends on the characteristics of the target representation. Future work could explore hybrid approaches integrating explicit linguistic knowledge with generative transformer models, potentially combining the interpretability of template-guided methods with the expressive capacity of large language models to improve performance in low-resource sign language translation scenarios.

## 6. Acknowledgments

The authors gratefully acknowledges the support of Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) through the research grant [CVU No. 1279958 and CVU No. 1288087], which made this study possible.

## Declaration on Generative AI

Grammarly was employed to suggest minor grammatical and stylistic corrections. No generative AI systems were used to write or edit the content of this manuscript.

## References


[1] Consejo Nacional para el Desarrollo y la Inclusión de las Personas con Discapacidad (CONADIS), Lengua de señas mexicana (lsm), https://www.gob.mx/conadis/articulos/lengua-de-senas-mexicana-lsm, 2016. Consultado: 24 mayo 2026.

[2] M. E. S. de Fleischmann, R. G. Pérez, Diccionario de Lengua de Señas, primera edición ed., Libre Acceso A.C. and Consejo Nacional para Prevenir la Discriminación (CONAPRED), 2011. URL: https://educacionespecial.sep.gob.mx/storage/recursos/2023/05/xzrfl019nV-4Diccionario_lengua_%20Senas.pdf.

[3] V. Burad, La glosa: Un sistema de notación para la lengua de señas, https://cultura-sorda.org/la-glosa-un-sistema-de-notacion-para-la-lengua-de-senas/, 2011. Cultura Sorda. Consultado: 24 mayo 2026.

[4] Instituto de las Personas con Discapacidad de la Ciudad de México, Glosario digital de lengua de señas mexicana, https://lsm.indiscapacidad.cdmx.gob.mx/, 2026. Consultado: 24 mayo 2026.

[5] H. Caballero Hernández, V. Muñoz Jiménez, M. A. Ramos Corchado, Desarrollo de software dedicado a la traducción de la lengua mexicana de señas mediante deep learning y machine learning, PAAKAT: Revista de Tecnología y Sociedad 15 (2025).

[6] A. Bonet-Jover, J. Á. González-Barba, L. Chiruzzo, Overview of IberLEF 2026: Natural Language Processing Challenges for Spanish and other Iberian Languages, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2026), co-located with the 42nd Conference of the Spanish Society for Natural Language Processing (SEPLN 2026), CEUR-WS. org, 2026.

[7] A. Y. Rodríguez-González, D. Fajardo-Delgado, M. Á. Álvarez-Carmona, M. G. Sánchez-Cervantes, Á. Díaz-Pacheco, R. Aranda, S. Fajardo-Flores, C. H. Martínez-Rodríguez, Overview of mslg-spa at iberlef 2026: Bidirectional translation between mexican sign language glosses and spanish, Procesamiento del Lenguaje Natural 77 (2026).

[8] Secretaría de Salud, Con discapacidad auditiva, 2.3 millones de personas: Instituto nacional de rehabilitación, https://www.gob.mx/salud/prensa/530-con-discapacidad-auditiva-2-3-millones-de-personas-instituto-nacional-de-rehabilitacion?idiom=es, 2022. Consultado: 24 mayo 2026.

[9] D. V. Lara-Ortiz, R. Q. Fuentes Aguilar, I. Chairez, Design of a neural transformer for spanish to mexican sign language automatic translation/interpretation, Network: Computation in Neural Systems 37 (2026) 206–232. URL: https://doi.org/10.1080/0954898X.2024.2435495. doi:10.1080/0954898X.2024.2435495.

[10] C. O. Sosa-Jiménez, H. V. Ríos-Figueroa, A. L. Solís-González-Cosío, A prototype for mexican sign language recognition and synthesis in support of a primary care physician, IEEE Access 10 (2022) 127620–127635. doi:10.1109/ACCESS.2022.3226696.

[11] V. Lara-Ortiz, R. Q. Fuentes-Aguilar, I. Chairez, Spanish to mexican sign language glosses corpus for natural language processing tasks, Scientific Data 12 (2025) 702. URL: https://doi.org/10.1038/s41597-025-04871-7. doi:10.1038/s41597-025-04871-7.

[12] V. Araujo, M. M. Trusca, R. Tufiño, M.-F. Moens, Sequence-to-sequence spanish pre-trained language models, 2023. arXiv:2309.11259.

[13] VaniLara, esp-to-lsm-barto-model, https://huggingface.co/VaniLara/esp-to-lsm-barto-model, 2024. Fine-tuned version of vgaraujov/bart-base-spanish for Spanish to Mexican Sign Language gloss translation.

[14] J. Cañete, G. Chaperon, R. Fuentes, J.-H. Ho, H. Kang, J. Pérez, Spanish pre-trained bert model and evaluation data, in: PML4DC at ICLR 2020, 2020.