人工知能学会第二種研究会資料
AIチャレンジ研究会
SIG-Challenge-063-05

Back Translation in Sign Language Generation

Khan Nabeela Khanum*[1], Tan Sihan[1], Itoyama Katsutoshi[1,2], and Nakadai Kazuhiro[1]

[1] Department of Systems and Control Engineering, School of Engineering, Tokyo Institute of Technology
[2]Honda Research Institute Japan Co.Ltd.

Abstract—This survey paper provides a comprehensive overview of the evolving landscape of Sign Language Back Translation (SLBT) for Sign Language Generation (SLG). BT is utilized to convert spoken text into gloss, a written form of SL, serving as a crucial intermediary step in SLG. The paper explores various SLBT paradigms, highlighting key studies and encapsulating SLBT's trajectory. It addresses challenges and promising advancements in back translation, aiming for seamless communication between spoken and signed languages.

1. INTRODUCTION

The application of sign languages (SLs) within the deaf community has become an established mode of communication. Sign languages utilize multiple channels, incorporating both manual and non-manual features. to break the linguistic and cultural barrier between hearing and deaf communities, translation plays an important role However, according to [1], manual translations is too expensive and a rather difficult to practice in daily life. In this case, the most affordable means of communication is using Machine Translation (MT)[2]. Unlike spoken languages, sign languages lack standardized written scripts, rendering advanced Machine Translation (MT) models designed for text-based languages unsuitable for SLs. To address this challenge, various writing methods have been introduced for SLs, including Glosses [3], SignWriting[4], HamNoSys[5], Stokoe Notations[6], and Si5s[7]. Among these methods, glosses have emerged as the most popular choice. Glosses involve labeling signs with words from the corresponding spoken language, often including affixes and markers, providing a bridge between the visual nature of sign languages and the written form of spoken languages.

Glosses play a vital role in various sign language (SL) processing tasks like recognition, translation, and generation. Despite their limitations in capturing the full linguistic richness of SL [8], they are crucial in Sign Language Translation (SLT) applications, facilitating communication between hearing and deaf communities, especially in education and interpretation. Additionally, glosses provide valuable parallel data for MT training in SL processing research. Researchers often employ glosses as intermediaries, translating between SL and spoken language. For instance, during SL to spoken language or vice versa translation, gloss-to-text or text-to-gloss conversion acts as an intermediary step. When generating SL from spoken language, the initial process involves converting text to glosses, which are then used for SL generation. This paper specifically focuses on the initial stage for SL generation, i.e converting text to glosses.

Several research initiatives in MT have been undertaken in the past for the back translation (BT) of SL in order to generate SL, showcasing the innovative efforts in the realm of sign language accessibility. One notable project, ViSiCAST, funded by the European Union, aimed to enhance accessibility for deaf citizens using virtual animation or avatars presenting sign language [9]. Another significant European venture, the eSIGN Project, contributed to this mission by developing avatar-based technology for American Sign Language. DePaul University played a pivotal role in this initiative, creating an avatar named 'Paula' capable of conveying all linguistic parameters of ASL while translating English to ASL [10]. Additionally, the ProDeaf project emerged as a pioneering effort, converting Portuguese text and voice into Portuguese Sign Language (LIBRAS) to facilitate seamless communication between the deaf and hearing communities [11].

In this paper, our focus is on the exploration of sign language back translation from text to gloss. The structure of the paper is organized as follows. Section 2 elucidates the fundamental distinctions between sign language and spoken language and presents a review of related works in machine translation for sign language back translation (SLBT) . Section 3 outlines the evaluation metrics employed in various machine translation approaches. Section 4 encompasses a discussion on the current state of research, including its limitations, and outlines potential future directions for further investigation.

2. RELATED WORKS

2·1 Sign language v/s Spoken language

SL and spoken language exhibit fundamental differences in lexicons, grammar rules, and structure. Contrary to common misconceptions, SL is not a universal language; rather, each country possesses its distinct SL, such as American Sign language , German Sign Language, Japanese Sign Language, Indian Sign Language, Chinese Sign Language, and Greek Sign Language [12]. In the realm of communication, SL relies on visual transmission, utilizing vision power instead of hearing power [13]. The mode of expression in SL involves intricate components, including single or both hands, facial expressions, and body movements. Notably, SL deviate from spoken languages in multiple aspects, ranging from grammatical variances to structural disparities and word order differences [14]. elucidate challenges in translating from Spanish to Spanish Sign Language, emphasizing issues like mapping semantic concepts to specific signs or generating multiple signs from one concept [15]. The translation process from Arabic text to Arabic Sign Language also faces hurdles due to grammatical rule discrepancies and differences in word order between the source and target languages [16]. Sequentiality is another significant distinction, where spoken languages follow a phonemic sequence, whereas SL incorporate non-sequential components concurrently, involving fingers, hands, and facial expressions [14]. This simultaneous nature is exemplified in Thai Sign Language, where the linguistic structure deviates from the linear organization of the Thai language [17]. Specific features unique to sign languages include non-manual components, the utilization of space as a lexical element, varied parts of speech represented by the same sign, the use of classifiers with morphological value, and diverse sentence structures [18]. These differences underscore the richness and complexity of sign languages, challenging preconceptions and highlighting the need for specialized linguistic understanding and translation methodologies.

2·2 Sign Language Back Translation (SLBT) methods

2·2.1 Rule based Machine translation (RBMT)

The rule-based approach to text-to-gloss translation, as discussed in [19], is an early method in machine translation. Rule-Based Machine Translation (RBMT) relies on predefined linguistic rules for both source and target languages. It involves morphological, syntactic, and semantic analyses in both languages. The RBMT system includes components such as source language morphological analysis, parsing, translation, target language morphological generation, and final parsing. The Vauquois Pyramid Fig. 1 illustrates the complexity of rule-based approaches, representing various linguistic levels. While structured, RBMT systems face challenges with idiomatic expressions but form the foundation for machine translation techniques in sign languages.

2·2.2 Corpus based Machine translation (CBMT)

Corpus-Based Machine Translation (CBMT) relies on bilingual text corpora for generating translations. While Rule-Based Machine Translation (RBMT) systems can produce accurate translations, they are

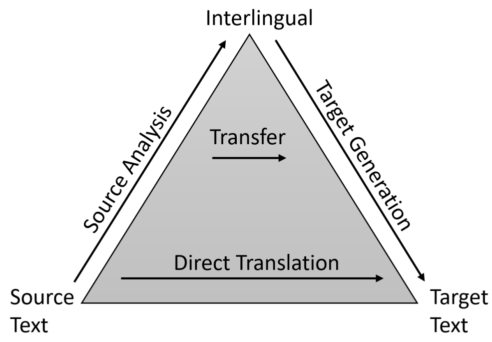

Fig. 1 Vauquois Pyramid

labor-intensive to develop, requiring manual crafting of linguistic resources and continuous rule additions, making the process time-consuming. In contrast, CBMT systems, also known as data-driven machine translations, leverage large bilingual datasets for translation.

Example based Machine Translation (EBMT)

Example-Based Machine Translation (EBMT), a concept first introduced by Makoto Nagao in 1984, relies on bilingual parallel corpora for its training, containing sentence pairs from both languages [20]. This approach has found pioneers in Sara Morrissey and Andy Way, who applied EBMT to Sign Language Machine Translation (SLMT) systems. Morrissey et al. employed the Marker Hypothesis to translate English to Dutch Sign Language, demonstrating its promise in segmenting English input text into chunks alignable with Sign Language annotations [21, 22]. Notably, ELAN annotation tool was utilized for sign language corpora, facilitating accurate alignment between English text and sign annotations. EBMT has also been successfully applied to languages with smaller corpora, such as Arabic Sign Language, where Almohimeed et al. utilized a 203-sentence corpus to translate Arabic text into Arabic Sign Language. The EBMT system operated on text chunks aligned with corresponding signs and, despite its limitations due to the quality of examples, yielded a word error rate (WER) of 46.7% and position-independent word error rate (PER) of 29.4% [23]. Additionally, researchers like Boulares et al. combined EBMT with genetic algorithms and fuzzy logic to translate English into American Sign Language (ASL). By integrating global and local alignment algorithms, they achieved effective proximity searches between words, demonstrating EBMT's potential for capturing complex linguistic structures [24]. For languages like Turkish Sign Language (TSL), where the grammar is poorly understood and datasets are limited, Selcuk-Simsek et al. proposed a bidirectional EBMT approach. Their system, incorporating a lexical supervision component (LSC) with morphological analyzers and disambiguation tools, achieved a BLEU score of 43% and a TER score of 38% using k-fold cross-validation [25]. Although EBMT excels in limited datasets, it faces challenges in scalability due to the need for a substantial number of high-quality examples. As detailed in Table 5, this approach's suitability diminishes for larger datasets, prompting a shift to explore the implementation of Statistical Machine Translation for such scenarios.

Statical Machine Translation (SMT)

Statistical Machine Translation (SMT) emerges as a significant player within the Corpus-Based Machine Translation (CBMT) paradigm. SMT, rooted in probability distributions and Bayesian approaches, operates efficiently with large bilingual corpora. Early pioneers such as Koehn et al. laid foundational work, exploring word alignment and various phrase translation methods [26, 27]. However, challenges surfaced, especially in smaller-scale applications like translating German text into German Sign Language (DGS) [28]. Researchers like [29] addressed these hurdles, employing morpho-syntactic analysis to enhance translation quality, leading to a notable 9% improvement [?]. Data scarcity posed a significant obstacle, prompting innovative solutions such as introducing thematic roles to capture verb meanings, as demonstrated in Chinese to Taiwanese Sign Language translation [30]. Syntactic-semantic information integration further enhanced translation accuracy, as seen in Spanish to LSE translation, where modules like categorization and Factored Translation Modules (FTMs) boosted BLEU scores significantly [31]. However, challenges persisted, exemplified in translating Indian Sign Language glosses, where existing models faced limitations [32]. Additionally, Turkish Sign Language (TID) translation efforts showcased diverse approaches, including stemming and semantic tagging, but lacked manual evaluation [30]. These endeavors underscore the ongoing pursuit to refine SMT for sign languages, emphasizing the need for innovative strategies and hybrid approaches to address data limitations and enhance translation outcomes.

Hybrid Machine Translation (HMT)

The integration of multiple machine translation systems within a single framework, known as Hybrid Machine Translation (HMT) systems, has emerged as a critical advancement. Addressing the limitations of single machine translation systems, researchers have explored diverse hybrid approaches, combining methodologies like example-based, transfer-based, knowledge-based, and statistical translation. For instance, Hogan et al. combined various translation sub-systems, showcasing the potential of hybridization [33]. Wu et al. adopted a hybrid model by merging rule-based and statistical approaches for translating Chinese to Taiwanese Sign Language, demonstrating substantial progress but highlighting challenges related to corpora size and extensibility [34]. Morrissey et al. contributed significantly to the field, utilizing MaTrEx Machine Translation system and integrating Statistical Machine Translation (SMT) and Example-Based Machine Translation (EBMT) methodologies. While achieving satisfactory results, challenges persisted in achieving natural sign language animations [35, 36]. San-Segundo et al. developed a comprehensive HMT approach incorporating rule-based, example-based, and statistical translators, enhancing the translation quality manifold. Their hierarchical structure and combination of techniques significantly improved results, outperforming individual techniques [37]. Additionally, researchers like Lopez-Ludena et al. refined hybrid systems, automating module generation and utilizing advanced translation strategies, producing remarkable improvements and paving the way for future advancements in SLMT [38, 39, 32]. These diverse hybridization efforts underscore the ongoing pursuit of enhancing SLMT, integrating advanced technologies, and addressing challenges, aiming for more accurate and natural sign language translations.

2·2.3 Neural Machine Translation (NMT)

Neural Machine Translation (NMT) stands as a transformative approach, utilizing artificial neural networks to predict word sequences. Manzano et al. employed NMT to translate English to American Sign Language (ASL) using the ASLG-PC dataset, yielding ASL glosses as output, albeit facing challenges due to a limited vocabulary size [40]. ATLASLang, translating Arabic to Arabic Sign Language, adopted NMT and surpassed previous systems with a BLEU score of 0.79 [41]. Text2Sign, an NMT system, utilized a Generative Adversarial Network and Motion Generation to produce sign videos from spoken language, showcasing robustness despite challenges related to avatar-based approaches [42]. Saunders et al. proposed an NMT approach focusing on automatic sign language production, resulting in enhanced Sign Language Production (SLP) performance [43]. Recognizing the importance of non-manual features, Saunders et al. extended their approach, encapsulating all sign articulators [44]. Ventura et al. advanced the field by generating realistic signing videos using the SIGNGAN approach, outperforming baseline systems both quantitatively and in human perception evaluations [45]. These studies highlight the evolving landscape of NMT in SLMT, underscoring both advancements and areas for further exploration

2·3 Datasets

For SLBT, a diverse array of datasets plays a pivotal role in advancing the field. These datasets serve as the foundation for training and evaluating models capable of translating sign language expressions back into spoken language. Among the prominent datasets, RWTH-PHOENIX-Weather-2014T (PHOENIX14T)

[46] stands out as a widely employed resource, offering German sign language videos, gloss, and spoken language text. This dataset, segmented into parallel sentences, has become a benchmark for evaluating baseline models in sign language translation (SLT). CSL-Daily [47], a notable dataset for Chinese SLT, covers a spectrum of themes with sign language videos featuring normative and natural expressions. RWTH-PHOENIX-Weather 2014 (PHOENIX14) focuses on German sign language videos sourced from weather news programs, providing valuable content for SLT endeavors. ASLG-PC12 [48], despite lacking sign language videos, presents a massive repository of gloss–text pairs, particularly suited for Gloss-to-Text tasks. Lastly, Spreadthesign-Ten (SP-10), a multilingual sign language recognition dataset, contributes to the broader understanding of sign languages by encompassing videos and corresponding texts from various linguistic backgrounds. These datasets collectively form a rich resource landscape, enabling researchers to explore and enhance sign language back translation techniques, ultimately fostering more inclusive and effective communication between signers and non-signers.

Notably, the release of How2Sign [49] in 2021 marked a significant stride. This multimodal and multi-view continuous American Sign Language (ASL) dataset introduces new dimensions to the field. While previous studies predominantly relied on the RWTH-PHOENIX-Weather-2014T (PHOENIX14T), How2Sign brings a fresh perspective by providing a broader and more comprehensive dataset for ASL. However, due to the dataset's complexity in data processing, researchers have yet to explore ASL production using deep learning techniques. The How2Sign dataset boasts a duration and vocabulary approximately 7.53 and 5.74 times larger than the commonly used PHOENIX14T, respectively. This significant expansion in both temporal coverage and vocabulary size holds promise for advancing the capabilities of deep learning models in handling the intricacies of ASL production, further enriching the resources available for sign language back translation research.

2·4 Evaluation Metrics

Accurate assessment of Sign Language Back Translation (SLBT) systems is crucial for determining their effectiveness. This evaluation encompasses both manual and automatic methods, each offering unique insights into the system's performance.Manual evaluations, drawing on feedback from deaf individuals and Sign Language (SL) experts, provide qualitative perspectives. They assess factors such as usability, appeal, and user satisfaction, offering a holistic understanding of the generated sign language output.

Automatic metrics, tailored to different translation types, include Word Error Rate (WER), Sentence Error Rate (SER), Position Independent Word Error Rate (PER), Translation Error Rate (TER), BLEU, and NIST [50, 51, 52, 53]. BLEU, a precision-based metric, shows a strong correlation with human judgment in machine translation [54].

Advancements in Neural Machine Translation (NMT) systems like ATLASLang and SIGNGAN introduce additional evaluation metrics, such as SSIM, PSNR, and MSE, providing nuanced assessments of synthesized SL images or videos [43, 42]. Complex NMT approaches may incorporate metrics like METEOR and RIBES to assess word order and reordering events [55].

Balancing manual and automatic evaluations is crucial, especially in Rule-Based Machine Translation (RBMT) scenarios with limited corpora. Ongoing research aims to refine and expand performance metrics, ensuring a comprehensive understanding of SLMT system outputs.

3. Discussion

The discussion on implications for research and practice in sign language back translation outlines promising avenues for future exploration, categorized into rule-based machine translation, corpus-based machine translation, neural machine translation, and sign generation.

In the realm of rule-based machine translation, addressing the translation of complex sentences and the lack of formal sign language grammar analysis are identified as critical research areas. The discussion underscores the importance of improving accuracy and usability, particularly in handling complex linguistic structures.

Moving to corpus-based machine translation, the challenges of limited bilingual corpora and the need for data acquisition in multiple sign languages emerge as key areas for future investigation. The efficient functioning of corpus-based systems is contingent on extensive datasets, prompting researchers to focus on creating diverse bilingual corpora for a comprehensive range of sign languages. Multilingual efforts, exemplified by existing datasets, like DICTA-SIGN and MultiATIS++ corpus, lay the foundation for future developments, aiming to bridge communication gaps within different deaf communities.

The discussion extends to neural machine translation, where the integration of deep learning and artificial intelligence (AI) is seen as a promising frontier. Notable successes, such as Google Translator's development based on neural machine translation (GNMT), point towards the potential of incorporating AI and deep learning strategies into prevalent systems. This approach holds the promise of achieving similar breakthroughs in text-to-sign translation across multiple languages.

Lastly, the importance of the sign generation system is emphasized, highlighting its crucial role in making more information and services available to the deaf community. Recognizing the significance of these avenues for future research, it becomes evident that addressing linguistic and cultural nuances, fostering interdisciplinary collaboration, and refining research methodologies are imperative. The discussion concludes by acknowledging the limitations of the review, notably the exclusion of articles in different languages and the study's focus solely on text-to-sign translation. These limitations underscore the ongoing need for inclusive and iterative approaches in advancing research in sign language back translation.

3·1 Limitations

Research on Sign Languages (SLs) reveals inherent challenges that pose obstacles to the development of robust linguistic models and technology. This section discusses two significant challenges: the Scarcity of SL Corpora and Ambiguity in Context.

3·1.1 Scarcity of SL corpora

The under-resourced nature of SLs, categorized as low-density languages, results in a scarcity of technological tools and computerized linguistic resources, such as corpora or lexicons. Corpora play a vital role in linguistic research, providing a corpus of naturally occurring signed language data for analysis and model training. The limited availability of SL corpora impedes the advancement of natural language processing in SLs, hindering the development of computational tools tailored to these languages.

3·1.2 Ambiguity in context

The absence of a standardized writing system for SLs and the reliance on video representations introduce challenges related to the ambiguity of context. Written languages rely on a standardized set of symbols, whereas SLs primarily use video for communication. This lack of a universally accepted writing system limits the creation and analysis of corpora, leading to challenges in disambiguating context in signed language utterances. In sign languages, the richness of facial expressions, body movements, and spatial components can introduce ambiguity in interpreting context. For instance, the same sign may have different meanings based on facial expressions or body language.

3·2 Future Direction

Efforts to surmount resource constraints in sign language back translation require a comprehensive strategy encompassing data augmentation, technological advancements, and collaborative endeavors. One key approach involves the expansion of sign language corpora through crowd-sourced initiatives, actively engaging the Deaf community. Collaborative platforms that facilitate the sharing of linguistic resources globally can foster a collective effort to alleviate resource scarcity, supporting the development of more inclusive machine translation models .

In addition to corpus expansion, researchers can explore transfer learning and pretraining models on larger, more general datasets to mitigate the impact of limited linguistic resources. Techniques such as domain adaptation and unsupervised learning can enhance model adaptation to the unique characteristics of sign language .

Complementing these strategies, the integration of data augmentation techniques proves instrumental in both overcoming resource constraints and enhancing translation accuracy. By synthetically expanding the training dataset through variations in signing speed, styles, facial expressions, and body movements, augmented datasets contribute to more robust models.

Moreover, future works should focus on developing context-aware models that account for the non-manual components, facial expressions, and body movements inherent in sign languages. This requires a nuanced understanding of the cultural and linguistic nuances embedded in sign language communication. So, a holistic approach to future research in SLBT should combine advancements in machine learning techniques, collaborative efforts, and an understanding of the unique linguistic aspects of sign languages. Integrating both general techniques for resource expansion and model adaptation, along with specific data augmentation methods, ensures a well-rounded strategy for addressing the challenges posed by resource constraints and enhancing the accuracy of sign language back translation.

4. Conclusion

In this paper, we survey the back translation in sign language for text to gloss conversion. We have discussed the importance of sign language back translation in sign language generation. we have also discussed the existing approaches and their limitations. We have explained the possible solutions of these limitations and some future directions with possible application of sign language back translation. In conclusion, this survey presents a comprehensive overview of sign language back translation, encapsulating key challenges and potential avenues for future research. The discussions span rule-based, corpus-based, and neural machine translation, each revealing distinct challenges and opportunities. The imperative to address the translation of complex sentences, formalize sign language grammar, and bridge data gaps in bilingual corpora for diverse sign languages underscores the interdisciplinary nature of this research. The potential breakthroughs promised by integrating deep learning and artificial intelligence into neural machine translation systems pose exciting prospects for advancing text-to-sign translation. The critical role of the sign generation system in fostering inclusivity and improving services for the deaf community is highlighted. Acknowledging limitations, the study advocates for ongoing refinement, inclusivity, and collaboration in shaping the trajectory of sign language back translation research.

References

[1] P. Isabelle and G. Foster, "Machine translation: overview," 2006.

[2] J. Porta, F. López-Colino, J. Tejedor, and J. Colás, "A rule-based translation from written spanish to spanish sign language glosses," *Computer Speech & Language*, vol. 28, no. 3, pp. 788–811, 2014.

[3] C. Valli and C. Lucas, *Linguistics of American sign language: An introduction.* Gallaudet University Press, 2000.

[4] A. C. da Rocha Costa and G. P. Dimuro, "Signwriting and swml: Paving the way to sign language processing," *Atelier Traitement Automatique des Langues des Signes, TALN*, vol. 2003, 2003.

[5] K. Kaur and P. Kumar, "Hamnosys to sigml conversion system for sign language automation," *Procedia Computer Science*, vol. 89, pp. 794–803, 2016.

[6] W. C. Stokoe Jr, "Sign language structure: An outline of the visual communication systems of the american deaf," *Journal of deaf studies and deaf education*, vol. 10, no. 1, pp. 3–37, 2005.

[7] R. A. Augustus, E. Ritchie, and S. Stecker, "The official american sign language writing textbook," *Los Angeles, CA: ASLized*, 2013.

[8] E. Pizzuto, P. Rossini, and T. Russo, "Representing signed languages in written form: questions that need to be posed," in *Proceedings of the Workshop on the Representation and Processing of Sign Languages, LREC*, vol. 2006, 2006.

[9] A. Othman and M. Jemni, "Designing high accuracy statistical machine translation for sign language using parallel corpus: case study english and american sign language," *Journal of Information Technology Research (JITR)*, vol. 12, no. 2, pp. 134–158, 2019.

[10] O. H. Al-Barahamtoshy and H. M. Al-Barhamtoshy, "Arabic text-to-sign (artts) model from automatic sr system," *Procedia Computer Science*, vol. 117, pp. 304–311, 2017.

[11] J. Rocha, J. Lensk, T. Ferreira, and M. Ferreira, "Towards a tool to translate brazilian sign language (libras) to brazilian portuguese and improve communication with deaf," in *2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pp. 1–4, IEEE, 2020.

[12] M. Alaghband, H. R. Maghroor, and I. Garibay, "A survey on sign language literature," *Machine Learning with Applications*, vol. 14, p. 100504, 2023.

[13] H.-Y. Su and C.-H. Wu, "Improving structural statistical machine translation for sign language with small corpus using thematic role templates as translation memory," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 7, pp. 1305–1315, 2009.

[14] V. Lopez Ludeña, R. San Segundo Hernández, C. González Morcillo, J. C. López López, and J. M. Pardo Muñoz, "Methodology for developing a speech into sign language translation system in a new semantic domain," 2012.

[15] R. San-Segundo, R. Barra, R. Córdoba, L. F. D'Haro, F. Fernández, J. Ferreiros, J. M. Lucas, J. Macías-Guarasa, J. M. Montero, and J. M. Pardo, "Speech to sign language translation system for spanish," *Speech Communication*, vol. 50, no. 11-12, pp. 1009–1020, 2008.

[16] H. Luqman and S. A. Mahmoud, "Automatic translation of arabic text-to-arabic sign language," *Universal Access in the Information Society*, vol. 18, pp. 939–951, 2019.

[17] S. Dangsaart, K. Naruedomkul, N. Cercone, and B. Sirinaovakul, "Intelligent thai text–thai sign translation for language learning," *Computers & Education*, vol. 51, no. 3, pp. 1125–1141, 2008.

[18] J. Porta, F. López-Colino, J. Tejedor, and J. Colás, "A rule-based translation from written spanish to spanish sign language glosses," *Computer Speech & Language*, vol. 28, no. 3, pp. 788–811, 2014.

[19] D. Kouremenos, K. Ntalianis, and S. Kollias, "A novel rule based machine translation scheme from greek to greek sign language: Production of different types of large corpora and language models evaluation," *Computer Speech & Language*, vol. 51, pp. 110–135, 2018.

[20] M. Nagao, "A framework of a mechanical translation between japanese and english by analogy principle," *Artificial and human intelligence*, pp. 351–354, 1984.

[21] N. Gough, *Example-based machine translation using the marker hypothesis.* PhD thesis, Dublin City University, 2005.

[22] S. Morrissey and A. Way, "An example-based approach to translating sign language," in *Workshop on example-based machine translation*, pp. 109–116, 2005.

[23] A. Almohimeed, M. Wald, and R. I. Damper, "Arabic text to arabic sign language translation system for the deaf and hearing-impaired community," in *Proceedings of the second workshop on speech and language processing for assistive technologies*, pp. 101–109, 2011.

[24] C. M. Bishop, "Mixture density networks," 1994.

[25] M. Selcuk-Simsek and I. Cicekli, "Bidirectional machine translation between turkish and turkish sign language: a data-driven approach," *Int. J. Nat. Lang. Comput*, vol. 6, no. 3, pp. 33–46, 2017.

[26] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, *et al.*, "Moses: Open source toolkit for statistical machine translation," in *Proceedings of the 45th annual meeting of the association for computational linguistics companion volume proceedings of the demo and poster sessions*, pp. 177–180, Association for Computational Linguistics, 2007.

[27] P. Koehn, F. J. Och, and D. Marcu, "Statistical phrase-based translation," in *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 127–133, 2003.

[28] J. Bungeroth and H. Ney, "Statistical sign language translation," in *sign-lang@ LREC 2004*, pp. 105–108, European Language Resources Association (ELRA), 2004.

[29] D. Stein, C. Schmidt, and H. Ney, "Analysis, preparation, and optimization of statistical sign language machine translation," *Machine Translation*, vol. 26, pp. 325–357, 2012.

[30] D. Stein, J. Bungeroth, and H. Ney, "Morpho-syntax based statistical methods for automatic sign language translation," in *Proceedings of the 11th Annual conference of the European Association for Machine Translation*, 2006.

[31] V. Lopez Ludeña, R. San Segundo Hernández, R. d. Córdoba Herralde, J. Ferreiros López, J. M. Montero Martínez, and J. M. Pardo Muñoz, "Factored translation models for improving a speech into sign language translation system," 2011.

[32] V. López-Ludeña, C. González-Morcillo, J. C. López, R. Barra-Chicote, R. Córdoba, and R. San-Segundo, "Translating bus information into sign language for deaf people," *Engineering Applications of Artificial Intelligence*, vol. 32, pp. 258–269, 2014.

[33] C. Hogan and R. E. Frederking, "An evaluation of the multi-engine mt architecture," in *Conference of the Association for Machine Translation in the Americas*, pp. 113–123, Springer, 1998.

[34] C.-H. Wu, H.-Y. Su, Y.-H. Chiu, and C.-H. Lin, "Transfer-based statistical translation of taiwanese sign language using pcfg," *ACM transactions on Asian language information processing (TALIP)*, vol. 6, no. 1, pp. 1–es, 2007.

[35] S. Morrissey, "Assistive translation technology for deaf people: translating into and animating irish sign language," 2008.

[36] S. Morrissey and A. Way, "Joining hands: Developing a sign language machine translation system with and for the deaf community," 2007.

[37] R. San-Segundo, J. M. Montero, R. Cordoba, V. Sama, F. Fernández, L. D'haro, V. López-Ludeña, D. Sánchez, and A. García, "Design, development and field evaluation of a spanish into sign language translation system," *Pattern Analysis and Applications*, vol. 15, pp. 203–224, 2012.

[38] V. López-Ludeña, R. San-Segundo, C. G. Morcillo, J. C. López, and J. M. P. Muñoz, "Increasing adaptability of a speech into sign language translation system," *Expert Systems with Applications*, vol. 40, no. 4, pp. 1312–1322, 2013.

[39] V. Lopez Ludeña, R. San Segundo Hernández, C. González Morcillo, J. C. López López, and J. M. Pardo Muñoz, "Methodology for developing a speech into sign language translation system in a new semantic domain," 2012.

[40] D. Manzano, "English to asl translator for speech2signs," *Retrieved March*, vol. 16, p. 2021, 2018.

[41] M. Brour and A. Benabbou, "Atlaslang mts 1: Arabic text language into arabic sign language machine translation system," *Procedia computer science*, vol. 148, pp. 236–245, 2019.

[42] S. Stoll, N. C. Camgoz, S. Hadfield, and R. Bowden, "Text2sign: towards sign language production using neural machine translation and generative adversarial networks," *International Journal of Computer Vision*, vol. 128, no. 4, pp. 891–908, 2020.

[43] B. Saunders, N. C. Camgoz, and R. Bowden, "Progressive transformers for end-to-end sign language production," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XI 16*, pp. 687–705, Springer, 2020.

[44] B. Saunders, N. C. Camgoz, and R. Bowden, "Adversarial training for multi-channel sign language production," *arXiv preprint arXiv:2008.12405*, 2020.

[45] B. Saunders, N. C. Camgoz, and R. Bowden, "Everybody sign now: Translating spoken language to photo realistic sign language video," *arXiv preprint arXiv:2011.09846*, 2020.

[46] N. C. Camgoz, S. Hadfield, O. Koller, H. Ney, and R. Bowden, "Neural sign language translation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7784–7793, 2018.

[47] H. Zhou, W. Zhou, W. Qi, J. Pu, and H. Li, "Improving sign language translation with monolingual data by sign back-translation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1316–1325, 2021.

[48] A. Othman and M. Jemni, "English-asl gloss parallel corpus 2012: Aslg-pc12," in *sign-lang@ LREC 2012*, pp. 151–154, European Language Resources Association (ELRA), 2012.

[49] A. Duarte, S. Palaskar, L. Ventura, D. Ghadiyaram, K. DeHaan, F. Metze, J. Torres, and X. Giro-i Nieto, "How2sign: a large-scale multimodal dataset for continuous american sign language," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2735–2744, 2021.

[50] A. Othman and M. Jemni, "English-asl gloss parallel corpus 2012: Aslg-pc12," in *sign-lang@ LREC 2012*, pp. 151–154, European Language Resources Association (ELRA), 2012.

[51] S. Morrissey and A. Way, "Lost in translation: the problems of using mainstream mt evaluation metrics for sign language translation," 2006.

[52] A. Kulesza and S. Shieber, "A learning approach to improving sentence-level mt evaluation," in *Proceedings of the 10th international conference on theoretical and methodological issues in machine translation*, European Association for Machine Translation, 2004.

[53] M. Snover, B. Dorr, R. Schwartz, L. Micciulla, and J. Makhoul, "A study of translation edit rate with targeted human annotation," in *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers*, pp. 223–231, 2006.

[54] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a method for automatic evaluation of machine translation," in *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pp. 311–318, 2002.

[55] H.-Y. Jung, J.-H. Lee, E. Min, and S.-H. Na, "Word reordering for translation into korean sign language using syntactically-guided classification," *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 19, no. 2, pp. 1–20, 2019.