Shedding Light on the Underexplored: Tackling the Minor Sign Language Research Topics

Jung-Ho Kim, Changyong Ko, Mathew Huerta-Enochian, and Seung Yong Ko
EQ4ALL
Nonhyeon-ro 76-gil 11, Gangnam-gu, Seoul, Republic of Korea
{stuartkim, ericko, mathew, stephenko}@eq4all.co.kr

Abstract

In the past decade, sign language research has achieved remarkable results alongside advancements in deep learning. However, there is a disconnect between the outcomes of these research efforts and the actual use of sign language by signers. In this position paper, we reviewed sign language papers related to deep learning published in the last ten years to explore the reasons for this gap. We found many areas of research that are still underdeveloped, despite their linguistic importance. Based on an analysis of known corpora and methodologies, we identified the reasons for the lack of progress in these areas and propose directions for future research efforts.

Keywords: sign language research, underexplored research topics, sign language linguistics, communication methodologies

1. Introduction

We have seen many advances in sign language research with the introduction of deep learning. The most significant advances have been in recognition (Rastgoo et al., 2021a), translation (Kahlon and Singh, 2023), and generation (Rastgoo et al., 2021b). Despite severely limited resources, sign language research continues to make new advances every year.

Nevertheless, there are elements of sign language that are not studied despite being important linguistic elements (eye-gaze, topic, role-shifting, tensions, space allocation, depicting signs, buoys, etc.). These are important aspects of the language that are used in real life and should be studied if we want to make the results of sign language research practical.

In this position paper, we examine the elements of sign language linguistics, and investigate both actively researched areas and those that have received less attention. Furthermore, we propose why such studies are significant, discuss why certain studies have not been well-conducted, and what actions we should take to facilitate research in these areas.

2. Sign Language Linguistics

Sign languages employ visual-manual modalities, involving handshapes, movements, facial expressions, and body postures to convey meaning (Sutton-Spence and Woll, 1999; Valli and Lucas, 2000; Sandler and Lillo-Martin, 2006). This distinct mode of communication leads to unique linguistic structures, including phonology, morphology, syntax, and semantics, tailored to the visual-spatial nature of sign languages. In this section, we delve into certain linguistic features that are more prominently highlighted in sign languages compared to spoken languages.

Space and simultaneity Sign languages are often referred to as spatial languages due to their inherent use of space to convey complex meanings. By exploiting the signing space with various articulators, signers can simultaneously present multiple pieces of information, a feature known as simultaneity (Geraci et al., 2008). For instance, signer use of buoys, which are handshapes or signs held in place to maintain a reference point or context while other signs are used to expand on other concepts, has been documented in various sign languages (Liddell, 2003; Tang et al., 2007). Simultaneous signs can represent actions, locations, or other descriptive information, allowing for a rich layering of language that is conveyed in a visually intuitive manner. This multi-layered approach to communication enables signers to present complex scenarios and narratives efficiently and effectively.

Topicalization Topicalization in sign languages involves the marking of a topic or the subject matter of a discourse at the beginning of a sentence or phrase, which is then followed by a comment or predicate about that topic. Friedman (1976); Aarons (1996) studied topicalization in American Sign Language (ASL) and Sze (2011) studied it in Hong Kong Sign Language (HKSL). This structure is often marked by specific non-manual signals such as raised eyebrows or a slight forward lean of the body, clearly distinguishing the topic from the rest of the discourse. This linguistic feature allows signers to structure their discourse in a way that

Proceedings of the 11th Workshop on the Representation and Processing of Sign Languages, pages 147–158

highlights the main points of discussion, making the communication clear and focused.

Role-shifting Role-shifting is a dynamic feature of sign languages where signers take on the roles of different characters in a narrative. Padden (1986) gave an early analysis of role-shifting in ASL and argued that it its use is more than just play-acting. By physically shifting their body orientation, facial expressions, and gaze, signers can represent different perspectives and viewpoints within a story. Role-shifting adds depth to a narrative by allowing the signer to embody different characters, making utterances more engaging and easier to follow. This technique is not only a powerful storytelling tool but also a sophisticated linguistic mechanism for indicating changes in subject, object, and possessive relationships within a narrative.

Phonology Sign language phonology encompasses both spatial and temporal aspects of signing, a notable difference from spoken language phonology. Brentari (1998) explored both the simultaneity of ASL phonemes and asserted that movements are the most basic prosodic elements of ASL. Brentari (2011) later presented a thorough overview of phonology in sign languages, focusing on ASL but also drawing from studies on other sign languages. Much research has been devoted to exploring the building blocks of signing across different sign languages, usually focusing on articulator position, orientation, shape, and movement in the signing space. Temporal phonological features such as prosody and rhythm are alos known to play a crucial role in most sign languages, adding layers of meaning and aiding in the conveyance of complex ideas and emotions. Cross-lingual variation has also been studied. For example, Tang et al. (2010) found that while eye blinks were used to mark certain intonational phrases in Japanese Sign Language (JSL), HKSL, Swiss German Sign Language (DSGS), and ASL, their use in JSL was unique out of these languages for blinks co-occurring with head nods rather than sign lengthening.

Non-lexical expressions Non-lexical expressions in sign languages encompass a range of communicative behaviors beyond the use of lexical signs, and include non-manual expressions (Valli and Lucas, 2000; Sandler and Lillo-Martin, 2006), depicting signs (Liddell, 2003; Cormier et al., 2012), and even gestures (Liddell and Metzger, 1998; Goldin-Meadow and Brentari, 2017). Non-manual expressions involve the use of facial expressions, body posture, and eye movements to convey meaning, mood, or grammatical information, adding depth and nuance to the signed message. Depicting signs use handshapes and movements to represent objects, actions, or concepts, often providing visual and spatial information about the subject matter. Gestures, although not strictly part of the formal sign language lexicon, are incorporated into communication, offering a universal means of conveying ideas or emotions, sometimes transcending linguistic boundaries. Together, these elements enrich the expressive capacity of sign languages, allowing for a dynamic and multifaceted mode of communication.

3. Sign Language Research Topics

We examined research topics in sign language studies that applied deep learning and selected several representative topics, as can be seen in Figure 1. We also identified research topics with relatively few or no publications, despite being important linguistic aspects of sign language.

3.1. Research Trends

We analyzed trends in sign language research from the past decade by reviewing a total of 544 papers from workshops, conferences, and journals in the fields of sign language, natural language processing, and computer vision. These papers were collected from the top twenty (by h5-index) publications in each of the following Google Scholar subcategories: Artificial Intelligence, Computational Linguistics, and Computer Vision & Pattern Recognition, in addition to selected papers from sign language workshops. The collection was restricted to works published between 2014 and early 2024. We categorized each paper by main topic and sub topic based on our interpretation of each paper's main focus. We provide our collection of relevant papers and paper topics through the digital repository link: https://doi.org/10.5281/zenodo.10948417.

Recognition Sign language recognition (SLR) involves automatically identifying handshapes, non-manual markers, fingerspellings, and glosses in video data and has seen the most active research (about 33% or 180 of 544 papers). Continuous and isolated SLRs are being advanced not only through improved feature extraction (He et al., 2016; Carreira and Zisserman, 2017; Xie et al., 2018) but also through new methods and applications such as better fusion of multiple input modalities (Chen et al., 2022), cross-frame feature trajectory analysis (Hu et al., 2023b), and knowledge distillation (Guo et al., 2023). Recently, sign spotting (Varol et al., 2022; Vázquez Enríquez et al., 2022) and sign segmentation (Woll et al., 2022; Moryossef et al., 2023) have

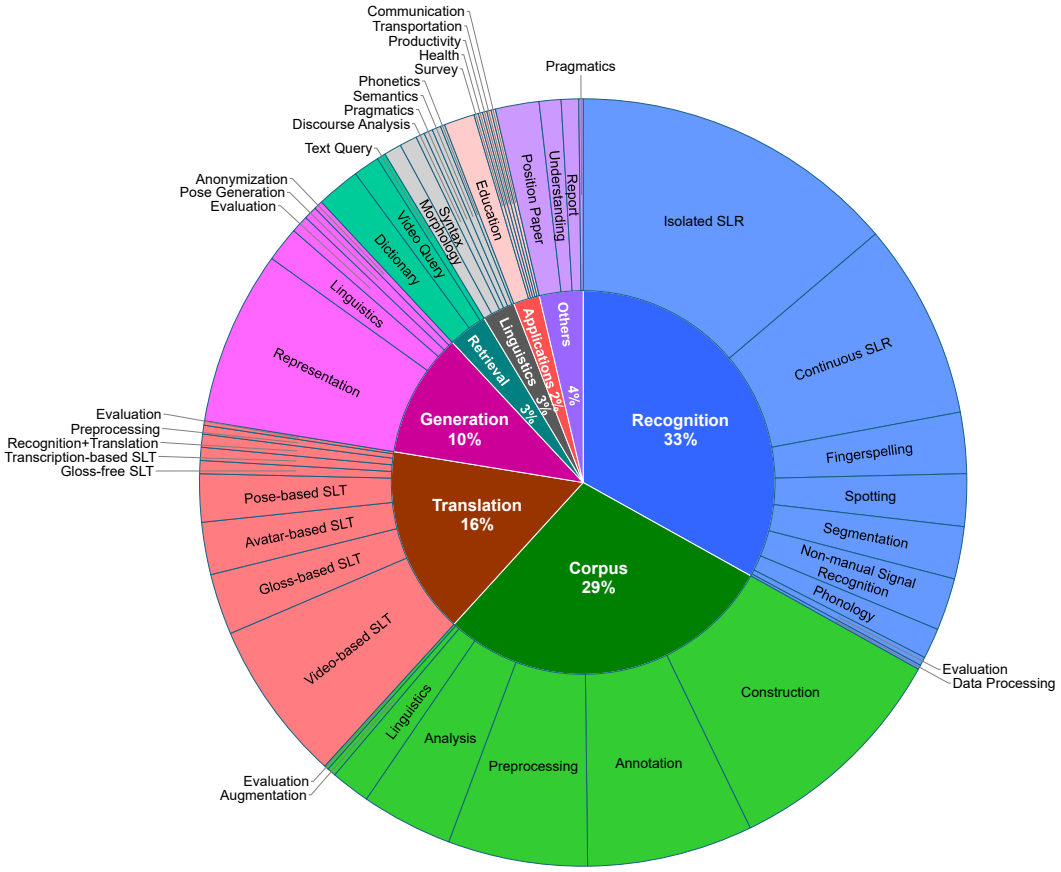

Figure 1: Research topics in sign languages. Tools for corpus, such as annotation and preprocessing, are included in each sub-topic depending on the purpose of the research.

emerged as prominant sub-topics.

Translation Sign language translation (SLT) is a task that translates between spoken language and sign language, or vice versa. Spoken language is represented through sound or text, and sign language is represented through gloss, skeletal pose, video (photo-realistic or avatar animation), or a notation system (usually SignWriting (Sutton, 2000) or HamNoSys (Hanke, 2004)). Recently, Gloss-free SLT, which translates sign language without the need for gloss supervision, has been actively researched (Yin et al., 2023; Lin et al., 2023; Zhou et al., 2023).

Generation Sign language generation (SLG) is the task of creating sign language poses or videos without translation[1]. Research has been conducted on a variety of topics, including the diversity of expressions (Kopf et al., 2023) and the anonymization of sign language users (Saunders et al., 2021; Xia et al., 2022). There have also been active proposals for research aimed at reflecting sign language linguistics in the generation process (McDonald et al., 2014).

Retrieval While research on sign language dictionaries supporting word-based or handshape-based search has been active for some time, recent studies have focused on information retrieval through natural language queries in text (Duarte et al., 2022; Cheng et al., 2023) or video data (Sedmidubsky et al., 2018; De Coster and Dambre, 2023).

Understanding Although less researched than the other main topics, sign language understanding (SLU) has been explored in several ways. Recent studies have proposed methods for linguistically modeling sign languages (Mocialov et al., 2018; Hu et al., 2023a). An interesting development is the proposal of research on coreference resolution (Yin et al., 2021a) and a call to recognize SLU as a field within natural language processing (Yin et al., 2021b).

Others Sign language corpora have been crucial linguistic assets in sign language research for

[1] In this paper, we classify approaches that include both translation and generation as SLT and approaches that involve only sign language generation as SLG.

an extended period, and corpora construction and analysis are areas that have received much focus. Additionally, applications and analysis of sign language in diverse areas such as health care, education, and communication have been proposed in academic papers. However, this paper focuses on deep learning-related research and does not extensively discuss these topics.

3.2. Underexplored Topics

Research in SLR, SLT, and SLG has advanced significantly, and yet some linguistic aspects of sign language modeling remain underexplored. This gap highlights a potential disconnect between research-generated output and actual signers' usage, underscoring the importance of incorporating sign language linguistics into future studies to ensure their authenticity and relevance. Below, we have listed research areas that, while being linguistically important in sign language, we believe are not being sufficiently researched.

Elicitation methodologies While elicitation methodologies have been studied extensively in traditional sign language corpora research and for spoken-language machine translation corpora, it has been mostly ignored in phrase-level sign language machine translation corpora, with few exceptions. Matthes et al. (2012) detailed how they developed tasks for capturing high-quality sign language utterances while still ensuring high overlap between multiple sign languages. Huerta-Enochian et al. (2022) compared several text-to-sign translation elicitation and revision methodologies and showed that text-based elicitation produced the least natural signing. Furthermore, we know that testing translation performance with back-translated data as the source language for spoken languages artificially inflates scores (Zhang and Toral, 2019; Graham et al., 2020), but bias in development methodologies have not yet been explored for SLT.

Pragmatics in SLT SLT has now reached a level of maturity where it is poised to explore practicality in additional to novelty. To enhance the practical use of SLT, it is necessary to contemplate how to deliver sign language expressions from a pragmatic perspective. For instance, when translating and generating sign language, space should be used in concert with non-manual signals in order to generate easily-understandable translations. Recognition systems should be designed to handle a wide range of signs and integrate naturally with users without needing special gloves, cameras, or lights. Recently, Fried et al. (2023) called for increased focus on pragmatics for large language models (LLMs), emphasizing the need for LLMs to adapt to the interlocutor. We suggest that this need is even greater for sign language modeling, given the crucial role of context in shaping how concepts are expressed.

Depicting signs Depicting signs are an area of research that is less frequently addressed in studies on SLR, SLT, and SLG. However, it is necessary to model depicting signs in each of these areas in order to approach the sign language representations actually used by signers. Since depicting signs are non-lexical expressions their use varies from person to person. There are many types of depicting signs, including the creation of gestures, the use of sign language to represent entities, and the description of situations through actions. Recent research on multi-modal large language models suggests new possibilities for exploring depicting signs. An important aspect of this research could be the representation of actions and relationships using one or both hands in sign language.

Rhythm and tension When generating sign language, the rhythm and stress of the signs are crucial elements in determining nuances. Similar to pragmatics, creating the appropriate sign language rhythm and stress according to the context will enable more natural sign language expressions and improve reception from the Deaf community.

Others There is a need for research on aspects that can be effectively used in sign language communication, such as topicalization and role-shifting. Moreover, translation between different sign languages could also present an intriguing area of study, potentially requiring methodologies distinct from those used in conventional translation. It is essential for research to more actively incorporate the history, culture, and linguistic aspects of sign language. There are also other areas in need of exploration, and we hope to see more proactive investigation of them in the future.

4. Challenges and Issues

We retrospectively examined existing studies with a focus on corpus and methodology challenges and explored how to resolve the issues identified.

4.1. Corpora

We examined a range of sign language corpora and summarized twenty-two commonly used corpora in Table 1. Here we argue that the following considerations should be taken into account in the use, management, and further construction of sign language corpora.

Corpus	Language	Access	Video	Size	Channel	License
ASLG-PC12 (Othman and Jemni, 2012)	ASL	open	N	77M (24M)	single	CC BY-NC 4.0
ATIS (Bungeroth et al., 2008)	multilingual (DGS,ISL,SASL)	restricted	Y	595 (–)	multi	CC BY-NC 4.0
AUSLAN (Johnston, 2009)	Auslan	restricted	Y	– (–)	multi	CC BY-NC-ND 4.0
BSL Corpus (Schembri et al., 2017)	BSL	open(partial) / academic	Y	– (14,754)	multi	custom
BOBSL (Albanie et al., 2021)	BSL	restricted	Y	1.2M (–)	multi	custom
CONTENT4ALL (SWISSTXT-WEATHER) (Camgöz et al., 2021)	DSGS	restricted	Y	811 (–)	single	CC BY-NC-SA 4.0
CONTENT4ALL (SWISSTXT-NEWS) (Camgöz et al., 2021)	DSGS	restricted	Y	6,031 (–)	single	CC BY-NC-SA 4.0
CONTENT4ALL (VRT-NEWS) (Camgöz et al., 2021)	VGT	restricted	Y	7,174 (–)	single	CC BY-NC-SA 4.0
Corpus NGT (Crasborn and Zwitserlood, 2008)	NGT	open(partial) / restricted	Y	– (490)	multi	CC BY-NC-SA 4.0
CSL-Daily (Zhou et al., 2021)	CSL	academic	Y	20,654 (–)	single	custom
Dicta Sign (Matthes et al., 2012)	BSL, DGS GSL, LSF	academic / restricted	Y	– (–)	single	–
KETI (Ko et al., 2019)	KSL	restricted	Y	2,940 (–)	single	–
KRSL-OnlineSchool (Mukushev et al., 2022)	KRSL	restricted	Y	1M (–)	single	–
NCSLGR (Neidle and Vogler, 2012)	ASL	open	Y	1,887 (1,874)	multi	custom
NIASL2021 (Huerta-Enochian et al., 2022)	KSL	open(domestic)	Y	201,026 (180,892)	multi	custom
DGS Corpus (Konrad et al., 2020)	DGS	open(partial) / restricted	Y	– (63,922)	multi	custom
RWTH-BOSTON-104 (Dreuw et al., 2007)	ASL	open	Y	201 (201)	single	–
RWTH-PHOENIX-WEATHER-2014-T (Camgoz et al., 2018)	DGS	open	Y	8,257 (8,257)	single	CC BY-NC-SA 3.0
SignBank♢	multilingual	open	N	– (29,035)	multi	–
STS Corpus (Öqvist et al., 2020)	SSL	open(web-access) / registered	Y	– (–)	multi	CC By-NC-SA 4.0
RWTH-PHOENIX-WEATHER 2014 (Forster et al., 2014)	DGS	open	Y	6,861 (6,841)	single	CC BY-NC-SA 3.0
How2Sign (Cardoso Duarte et al., 2021)	ASL	open(w/o gloss)	Y	35,191 (–)	single	CC BY-NC 4.0

♢: https://www.signbank.org/signpuddle/, accessed on February 23, 2024

Table 1: Summary of reviewed corpora. We limited reporting to *sentence-level* data. **Access**: *open*, *registered* (available with registration), *academic* (available for non-commercial research or academia), and *restricted* (available only with explicit permission). We report multiple levels as applicable. **Size**: The reported sentence-level instance count and our calculated open access count, if available. Every effort was made to report correct sizes for open access data, but there may be some deviation based on access method. **Channel**: Data is categorized based on the presence of annotations for separate hands or for non-manual signals, regardless of the existence of multiple tiers. **License**: The current corpus license. Note that licenses may differ from those reported in original research or from software licenses.

Data format The central challenge to choosing a data annotation format is that sign representation fidelity is inversely related to representation simplicity. In other words, simple representations like glosses cannot adequately represent the nuances of multiple signed instances while more informative representations like sign writing or even pose data are not easy to work with. This leads to variations in data and glossing formats across corpora, which in turn requires significant additional preprocessing before corpora can be used for training (De Sisto et al., 2022). Recently, there has been more inter-

est in rectifying this issue as can be seen in the proposed rectification of annotations from the easier project Kopf et al. (2022) and in Schulder et al. (2023) proposal of the sign language interchange format. While rectifying these differences between corpora is a good and necessary solution, using more unified annotation conventions for future corpus projects will be immensely helpful.

Data availability Though many corpora have been released for sign language research, collection and use of potential corpora is complicated by missing data links, mixed access levels, and custom licenses. Notably, some corpora were publicly available at the time of publication but are no longer accessible.

Commercial-friendly data Only two of the corpora we reviewed explicitly support commercial applications: the partially open release of the BSL Corpus and NIASL2021 (which is currently limited to users in Korea). In addition, five of the corpora do not include specific licensing information, introducing legal risks if used. The vast majority of corpora use derivatives of CC BY-NC or custom licenses that designate corpora for research purposes only. To encourage research from industry as well as academia, it may be necessary to reflect an incentive mechanism for data disclosure. However, in this case, ethical considerations such as re-obtaining consent from contributors due to a changing release policy and data anonymization should also be taken into account.

Signing quality Sign language corpora for machine learning show much variation in terms of signing quality. One major factor in this is the range of elicitation and collection methodologies. Some corpora feature only spontaneous utterances on open-ended topics, some corpora focus elicitation to specific tasks, and many corpora use either real-time interpretation or pre-translated utterances. We are not advocating against using specific corpora. On the contrary, given the small size of available data, utilizing existing corpora as much as possible—including corpora containing non-spontaneous signing—is necessary. While the effectiveness of high-quality training data is undisputed, lower-quality data is often utilized for pretraining, contrastive training, and other novel approaches. A key challenge moving forward will be to better classify signing data by recommended use and to improve elicitation techniques in general.

4.2. Methodologies

Text-to-sign translation There is a growing interest in direct pose- and video-predicting models, likely due to the lower annotation burden and the appeal of end-to-end solutions. While visualization of single-channel gloss data is limited, it still has significant value for identifying bias, data balance issues, linguistic insights, for researchers invested in procedural generation, and in hybrid approaches. Similarly, high-cost annotations like multi-channel glosses and notation systems offer the possibility of higher fidelity translations in specific domains. While we agree that the potential of end-to-end solutions are the most promising in the long-term, we urge the community to keep prioritizing multiple data modalities given the continued need for both short-term and long-term solutions.

Modeling non-lexical signs Procedural generation of non-lexical signs from gloss annotations is extremely challenging. High-detail annotations like multi-channel glosses, AZee, HamNoSys, SignWriting, and other phonetic annotations provide additional possibilities for non-lexical sign generation. While end-to-end solutions should be able to produce non-lexical signs, hybrid approaches like the one explored by Saunders et al. (2022) are likely more realistic in the short-term.

Non-lexical sign recognition is also an area that may likely benefit from novel approaches, particularly by delving into the intricacies of sign language. Effective recognition of non-lexical signs may involve understanding sign language morphemes, identifying what entities the handshapes represent, or even interpreting the intent behind gestures. This deeper comprehension could lead to more effective communication aids for the deaf and hard of hearing, by not just recognizing signs as whole units but understanding their component parts and the meanings they convey in different contexts.

SLT automatic evaluation While traditional machine translation metrics (BLEU, Rouge, etc.) can be applied to simple gloss translations, there is no definitive metrics for the many other text→sign output representations. Recently applied and proposed metrics include SignBLEU (Kim et al., 2024) for multi-channel gloss prediction; BLEU, chrF2++, and mean absolute error metrics for Formal SignWriting ASCII (Jiang et al., 2023); and SLR pose classification accuracy (Xiao et al., 2020) and Frechét Gesture Distance (Yoon et al., 2020) for pose prediction. As a community, we need to continue researching and improving potential metrics.

Human evaluation accessibility A hugely influential factor in the creation, evaluation, and curation of text-based data has been Amazon's Mechanical Turk (MTurk). While MTurk can be used for sign language data, finding highly-specialized participants through MTurk is a known challenge (Chandler and Shapiro, 2016). Furthermore, due to

communication barriers and ethics board requirements, most human evaluation of sign language is conducted locally and limited to the local sign language. This means that most machine learning research is human-validated on a single sign language or none at all. We see the need for better international cooperation, preferably as an official network, devoted to ensuring high-quality human evaluation for sign language applications.

5. Possible Directions

In the previous section, we placed as much importance on the practicality of sign language research outcomes as on improving performance. Here, we provide insights into areas that we think should receive increased focus in future research.

Additional annotation of existing corpora We have observed that understanding the position and direction in sign language plays a crucial role in comprehending its syntax. Therefore, transcribing this information, either automatically or manually, and applying it to SLR, SLT, and SLG models is essential.

Elicitation methodologies In light of the data scarcity problem in SLR, SLT, and SLG, the quality of signing data is of increased importance, and there are several urgent research directions to be explored. Data for specific translation applications usually requires highly-structured translations from existing spoken-language text. However, improving the quality of text-to-sign translations while ensuring high content fidelity is an open problem. As mentioned in section 3, traditional corpora research suggests using language-neutral elicitation materials, but applying such media to translation of specific phrases needs more exploration. In order to avoid bias, we need to research proper methodologies for sign language translation train and test set construction.

Pragmatics Pragmatics in sign language explores how language functions within social contexts and interactions. To address this, a deep learning model methodology is essential—one that not only minimizes ambiguity but also ensures communication objectives are met through word choice, spatial utilization, and the use of non-manual expressions. Establishing a clear evaluation framework is equally crucial to assess a model's overall effectiveness in enhancing clarity and communication efficiency. Fried et al. (2023) proposed how to model pragmatics with large language models to achieve these communication goals for all natural languages.

Non-lexical signs Effective modeling of non-lexical signs will require novel solutions, and we expect that many potential solutions will be found in linguistic insights. For example, Taub (2001) first proposed the analogue-building model process which is comprised of three steps (image selection, schematization, and encoding), and subsequent studies (Emmorey, 2014; Nordheimer et al., 2024) have built on and applied this model. We see the potential of this method applied to SLT through an approach using knowledge distillation and representation learning as a way to train entity translation in a generalizable way.

Hate speech The exploration of hate speech in sign language research is essential for the development of protective measures and educational tools that can help safeguard communities from discrimination and abuse. The nuanced gestures and expressions unique to sign languages can convey complex emotions and intentions, making it vital to understand how hate speech manifests in these modes of communication. Consequently, building comprehensive corpora that capture the breadth of sign language expressions, including those that could be considered hate speech, is imperative. These corpora will not only facilitate the identification and mitigation of hate speech within sign language communication but also contribute to the broader efforts of promoting digital safety and inclusiveness for all, regardless of mode of communication.

Deaf involvement Currently, Deaf involvement in sign language machine learning research is largely limited to participation in corpora construction and annotation and in human evaluation of developing technologies. Limited Deaf involvement in research means that hearing-centric views may grow unchecked and we risk losing sight of meaningful research objectives. On the other hand, increased involvement will provide insights to which non-native signers do not have access and ensure that we work towards developing solutions that the community can actually use.

6. Concluding Remarks

We have explored areas within sign language research that have not been well addressed. We also examined and proposed directions for future research in these areas. We argue that future sign language studies should be more closely connected with sign language linguistics and reconsider their practicality. We hope that by doing so, research outcomes will be more readily accepted in Deaf communities. Not all research topics could be covered in this paper, and as research progresses,

downstream tasks of NLP that are currently underexplored for sign language, including summarization, question answering, and language modeling, will likely receive more attention.

Acknowledgment

This work was partly supported by the Institute for Information and communications Technology Promotion (IITP) grant funded by the Korean government (MSIT) (No. 2022-0-00010, Development of Korean sign language translation service technology for the deaf in medical environment) and Artificial intelligence industrial convergence cluster development project funded by the Korean government (MSIT) & Gwangju Metropolitan City (No. BA00000797, LLM-based sign language translation for weather forecasts).

Sign Language Abbreviations

ASL American Sign Language
Auslan Australian Sign Language
BSL British Sign Language
CSL Chinese Sign Language
DGS German Sign Language
DSGS Swiss-German Sign Language
GSL Greek Sign Language
ISL Irish Sign Language
KRSL Kazakh–Russian Sign Language
KSL Korean Sign Language
LSF French Sign Language
NGT Dutch Sign Language
SASL South African Sign Language
SSL Swedish Sign Language
VGT Flemish Sign Language

Bibliographical References

Debra Aarons. 1996. Topics and topicalization in american sign language. *Stellenbosch Papers in Linguistics*, 30(1):65–106.

Samuel Albanie, Gül Varol, Liliane Momeni, Hannah Bull, Triantafyllos Afouras, Himel Chowdhury, Neil Fox, Bencie Woll, Rob Cooper, Andrew McParland, et al. 2021. Bbc-oxford british sign language dataset. *arXiv preprint arXiv:2111.03635*.

Diane Brentari. 1998. *A prosodic model of sign language phonology*. Mit Press.

Diane Brentari. 2011. Sign language phonology. *The handbook of phonological theory*, pages 691–721.

Jan Bungeroth, Daniel Stein, Philippe Dreuw, Hermann Ney, Sara Morrissey, Andy Way, and Lynette van Zijl. 2008. The ATIS sign language corpus. In *6th International Conference on Language Resources and Evaluation (LREC 2008)*, pages 2943–2946, Marrakech, Morocco. European Language Resources Association (ELRA).

Necati Cihan Camgoz, Simon Hadfield, Oscar Koller, Hermann Ney, and Richard Bowden. 2018. Neural sign language translation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7784–7793.

Necati Cihan Camgöz, Ben Saunders, Guillaume Rochette, Marco Giovanelli, Giacomo Inches, Robin Nachtrab-Ribback, and Richard Bowden. 2021. Content4all open research sign language translation datasets. In *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*, pages 1–5. IEEE.

Amanda Cardoso Duarte, Shruti Palaskar, Lucas Ventura Ripol, Deepti Ghadiyaram, Kenneth DeHaan, Florian Metze, Jordi Torres Viñals, and Xavier Giró Nieto. 2021. How2sign: A large-scale multimodal dataset for continuous american sign language. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition: Virtual, 19-25 June 2021: proceedings*, pages 2734–2743. Institute of Electrical and Electronics Engineers (IEEE).

Joao Carreira and Andrew Zisserman. 2017. Quo vadis, action recognition? a new model and the kinetics dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Jesse Chandler and Danielle Shapiro. 2016. Conducting clinical research using crowdsourced convenience samples. *Annual review of clinical psychology*, 12:53–81.

Yutong Chen, Ronglai Zuo, Fangyun Wei, Yu Wu, Shujie Liu, and Brian Mak. 2022. Two-stream network for sign language recognition and translation. *Advances in Neural Information Processing Systems*, 35:17043–17056.

Yiting Cheng, Fangyun Wei, Jianmin Bao, Dong Chen, and Wenqiang Zhang. 2023. Cico: Domain-aware sign language retrieval via cross-lingual contrastive learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 19016–19026.

Kearsy Cormier, David Quinto-Pozos, Zed Sevcikova, and Adam Schembri. 2012. Lexicalisation and de-lexicalisation processes in sign languages: Comparing depicting constructions and

viewpoint gestures. *Language & communication*, 32(4):329–348.

Onno Crasborn and Inge Zwitserlood. 2008. The Corpus NGT: an online corpus for professionals and laymen. In *Proceedings of the LREC2008 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora*, pages 44–49, Marrakech, Morocco. European Language Resources Association (ELRA).

Mathieu De Coster and Joni Dambre. 2023. Querying a sign language dictionary with videos using dense vector search. In *Proceedings of the 8th International Workshop on Sign Language Translation and Avatar Technology*. IEEE.

Mirella De Sisto, Vincent Vandeghinste, Santiago Egea Gómez, Mathieu De Coster, Dimitar Shterionov, and Horacio Saggion. 2022. Challenges with sign language datasets for sign language recognition and translation. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 2478–2487, Marseille, France. European Language Resources Association.

Philippe Dreuw, David Rybach, Thomas Deselaers, Morteza Zahedi, and Hermann Ney. 2007. Speech recognition techniques for a sign language recognition system. *hand*, 60:80.

Amanda Duarte, Samuel Albanie, Xavier Giró-i Nieto, and Gül Varol. 2022. Sign language video retrieval with free-form textual queries. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14094–14104.

Karen Emmorey. 2014. Iconicity as structure mapping. *Philosophical transactions of the Royal Society B: Biological sciences*, 369(1651):20130301.

Jens Forster, Christoph Schmidt, Oscar Koller, Martin Bellgardt, and Hermann Ney. 2014. Extensions of the sign language recognition and translation corpus rwth-phoenix-weather. In *LREC*, pages 1911–1916.

Daniel Fried, Nicholas Tomlin, Jennifer Hu, Roma Patel, and Aida Nematzadeh. 2023. Pragmatics in language grounding: Phenomena, tasks, and modeling approaches. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 12619–12640, Singapore. Association for Computational Linguistics.

Lynn A Friedman. 1976. The manifestation of subject, object, and topic in american sign language. *Word Order and Word Order Change*, pages 940–961.

Carlo Geraci, Marta Gozzi, Costanza Papagno, and Carlo Cecchetto. 2008. How grammar can cope with limited short-term memory: Simultaneity and seriality in sign languages. *Cognition*, 106(2):780–804.

Susan Goldin-Meadow and Diane Brentari. 2017. Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and brain sciences*, 40:e46.

Yvette Graham, Barry Haddow, and Philipp Koehn. 2020. Statistical power and translationese in machine translation evaluation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 72–81, Online. Association for Computational Linguistics.

Leming Guo, Wanli Xue, Qing Guo, Bo Liu, Kaihua Zhang, Tiantian Yuan, and Shengyong Chen. 2023. Distilling cross-temporal contexts for continuous sign language recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10771–10780.

Thomas Hanke. 2004. Hamnosys-representing sign language data in language resources and language processing contexts. In *LREC*, volume 4, pages 1–6.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Hezhen Hu, Weichao Zhao, Wengang Zhou, and Houqiang Li. 2023a. Signbert+: Hand-model-aware self-supervised pre-training for sign language understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9):11221–11239.

Lianyu Hu, Liqing Gao, Zekang Liu, and Wei Feng. 2023b. Continuous sign language recognition with correlation network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2529–2539.

Mathew Huerta-Enochian, Du Hui Lee, Hye Jin Myung, Kang Suk Byun, and Jun Woo Lee. 2022. Kosign sign language translation project: Introducing the niasl2021 dataset. In *Proceedings of the 7th International Workshop on Sign Language Translation and Avatar Technology: The Junction of the Visual and the Textual: Challenges and Perspectives*, pages 59–66.

Zifan Jiang, Amit Moryossef, Mathias Müller, and Sarah Ebling. 2023. Machine translation between spoken languages and signed languages represented in SignWriting. In *Findings of the Association for Computational Linguistics: EACL 2023*, pages 1706–1724, Dubrovnik, Croatia. Association for Computational Linguistics.

Trevor Johnston. 2009. Creating a corpus of auslan within an australian national corpus. In *Selected Proceedings of the 2008 HCSNet Workshop on Designing the Australian National Corpus: Mustering Languages*.

Navroz Kaur Kahlon and Williamjeet Singh. 2023. Machine translation from text to sign language: a systematic review. *Universal Access in the Information Society*, 22(1):1–35.

Jung-Ho Kim, Mathew Huerta-Enochian, Changyong Ko, and Du Hui Lee. 2024. Signbleu: Automatic evaluation of multi-channel sign language translation. Accepted.

Sang-Ki Ko, Chang Jo Kim, Hyedong Jung, and Choongsang Cho. 2019. Neural sign language translation based on human keypoint estimation. *Applied sciences*, 9(13):2683.

Reiner Konrad, Thomas Hanke, Gabriele Langer, Dolly Blanck, Julian Bleicken, Ilona Hofmann, Olga Jeziorski, Lutz König, Susanne König, Rie Nishio, Anja Regen, Uta Salden, Sven Wagner, Satu Worseck, Oliver Böse, Elena Jahn, and Marc Schulder. 2020. Meine dgs – annotiert. öffentliches korpus der deutschen gebärdensprache, 3. release / my dgs – annotated. public corpus of german sign language, 3rd release.

Maria Kopf, Rehana Omardeen, and Davy Van Landuyt. 2023. Representation matters: The case for diversifying sign language avatars. In *Proceedings of the 8th International Workshop on Sign Language Translation and Avatar Technology*. IEEE.

Maria Kopf, Marc Schulder, Thomas Hanke, and Sam Bigeard. 2022. Specification for the harmonization of sign language annotations. Project Deliverable EASIER D6.2, EASIER project, IDGS, Hamburg University, Hamburg, Germany.

Scott K Liddell. 2003. *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press.

Scott K Liddell and Melanie Metzger. 1998. Gesture in sign language discourse. *Journal of pragmatics*, 30(6):657–697.

Kezhou Lin, Xiaohan Wang, Linchao Zhu, Ke Sun, Bang Zhang, and Yi Yang. 2023. Gloss-free end-to-end sign language translation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 12904–12916, Toronto, Canada. Association for Computational Linguistics.

Silke Matthes, Thomas Hanke, Anja Regen, Jakob Storz, Satu Worseck, Eleni Efthimiou, Athanasia-Lida Dimou, Annelies Braffort, John Glauert, and Eva Safar. 2012. Dicta-sign-building a multilingual sign language corpus. In *5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon. Satellite Workshop to the eighth International Conference on Language Resources and Evaluation (LREC-2012)*.

John C. McDonald, Rosalee Wolfe, Robyn Moncrief, and Souad Baowidan. 2014. Analysis for synthesis: Investigating corpora for supporting the automatic generation of role shift. In *Proceedings of the LREC2014 6th Workshop on the Representation and Processing of Sign Languages: Beyond the Manual Channel*, pages 117–122, Reykjavik, Iceland. European Language Resources Association (ELRA).

Boris Mocialov, Helen Hastie, and Graham Turner. 2018. Transfer learning for British Sign Language modelling. In *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)*, pages 101–110, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Amit Moryossef, Zifan Jiang, Mathias Müller, Sarah Ebling, and Yoav Goldberg. 2023. Linguistically motivated sign language segmentation. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 12703–12724, Singapore. Association for Computational Linguistics.

Amit Moryossef, Kayo Yin, Graham Neubig, and Yoav Goldberg. 2021. Data augmentation for sign language gloss translation. In *Proceedings of the 1st International Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL)*, pages 1–11, Virtual. Association for Machine Translation in the Americas.

Medet Mukushev, Aigerim Kydyrbekova, Vadim Kimmelman, and Anara Sandygulova. 2022. Towards large vocabulary kazakh-russian sign language dataset: Krsl-onlineschool. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources*, pages 154–158.

Carol Neidle and Christian Vogler. 2012. A new web interface to facilitate access to corpora: Development of the asllrp data access interface (dai). In *Proc. 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon, LREC*, volume 3. Citeseer.

Swetlana Nordheimer, Allison Marlow, and Janina Scholtz. 2024. Fostering mathematical creativity and talents with mathematical problems and competitions in german sign language. *The 13th IMCGC Bloemfontein*.

Zrajm Öqvist, Nikolaus Riemer Kankkonen, and Johanna Mesch. 2020. STS-korpus: A sign language web corpus tool for teaching and public use. In *Proceedings of the LREC2020 9th Workshop on the Representation and Processing of Sign Languages: Sign Language Resources in the Service of the Language Community, Technological Challenges and Application Perspectives*, pages 177–180, Marseille, France. European Language Resources Association (ELRA).

Achraf Othman and Mohamed Jemni. 2012. English-asl gloss parallel corpus 2012: Aslg-pc12. In *sign-lang@ LREC 2012*, pages 151–154. European Language Resources Association (ELRA).

Carol Padden. 1986. Verbs and role-shifting in american sign language. In *Proceedings of the fourth national symposium on sign language research and teaching*, volume 44, page 57. National Association of the Deaf Silver Spring, MD.

Pierre Poitier, Jérôme Fink, and Benoît Frénay. 2024. Towards better transition modeling in recurrent neural networks: The case of sign language tokenization. *Neurocomputing*, 567:127018.

Razieh Rastgoo, Kourosh Kiani, and Sergio Escalera. 2021a. Sign language recognition: A deep survey. *Expert Systems with Applications*, 164:113794.

Razieh Rastgoo, Kourosh Kiani, Sergio Escalera, and Mohammad Sabokrou. 2021b. Sign language production: A review. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3451–3461.

Wendy Sandler and Diane Carolyn Lillo-Martin. 2006. *Sign language and linguistic universals*. Cambridge University Press.

Ben Saunders, Necati Cihan Camgoz, and Richard Bowden. 2020. 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*, pages 687–705. Springer.

Ben Saunders, Necati Cihan Camgoz, and Richard Bowden. 2021. Anonysign: Novel human appearance synthesis for sign language video anonymisation. In *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*, pages 1–8.

Ben Saunders, Necati Cihan Camgoz, and Richard Bowden. 2022. Signing at scale: Learning to co-articulate signs for large-scale photo-realistic sign language production. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5141–5151.

Adam C. Schembri, Jordan B Fenlon, Ramas Rentelis, and Kearsy Cormier. 2017. British sign language corpus project: A corpus of digital video data and annotations of british sign language 2008-2017.

Marc Schulder, Sam Bigeard, Thomas Hanke, and Maria Kopf. 2023. The sign language interchange format: Harmonising sign language datasets for computational processing. In *2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops: Sign Language Translation and Avatar Technology (SLTAT 2023)*, Rhodes, Greece. IEEE.

Jan Sedmidubsky, Petr Elias, and Pavel Zezula. 2018. Effective and efficient similarity searching in motion capture data. *Multimedia Tools and Applications*, 77:12073–12094.

Valerie Sutton. 2000. Signwriting. *Deaf Action Committee (DAC) for Sign Writing*.

Rachel Sutton-Spence and Bencie Woll. 1999. *The linguistics of British Sign Language: an introduction*. Cambridge University Press.

Felix Sze. 2011. Nonmanual markings for topic constructions in hong kong sign language. *Sign Language & Linguistics*, 14(1):115–147.

Gladys Tang, Diane Brentari, Carolina González, and Felix Sze. 2010. *Crosslinguistic variation in prosodic cues*. na.

Gladys Tang, Felix Sze, Scholastica Lam, et al. 2007. Acquisition of simultaneous constructions by deaf children of hong kong sign language. *Simultaneity in signed languages*, pages 283–316.

Sarah F Taub. 2001. *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge University Press.

Clayton Valli and Ceil Lucas. 2000. *Linguistics of American sign language: An introduction*. Gallaudet University Press.

Gül Varol, Liliane Momeni, Samuel Albanie, Triantafyllos Afouras, and Andrew Zisserman. 2022. Scaling up sign spotting through sign language dictionaries. *International Journal of Computer Vision*, 130(6):1416–1439.

Manuel Vázquez Enríquez, José L Alba Castro, Laura Docio Fernandez, Julio CS Jacques Junior, and Sergio Escalera. 2022. Eccv 2022 sign spotting challenge: Dataset, design and results. In *European Conference on Computer Vision*, pages 225–242.

Bencie Woll, Neil Fox, and Kearsy Cormier. 2022. Segmentation of signs for research purposes: Comparing humans and machines. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources*, pages 198–201, Marseille, France. European Language Resources Association (ELRA).

Zhaoyang Xia, Yuxiao Chen, Qilong Zhangli, Matt Huenerfauth, Carol Neidle, and Dimitris Metaxas. 2022. Sign language video anonymization. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources*, pages 202–211, Marseille, France. European Language Resources Association (ELRA).

Qinkun Xiao, Minying Qin, and Yuting Yin. 2020. Skeleton-based chinese sign language recognition and generation for bidirectional communication between deaf and hearing people. *Neural networks*, 125:41–55.

Saining Xie, Chen Sun, Jonathan Huang, Zhuowen Tu, and Kevin Murphy. 2018. Rethinking spatiotemporal feature learning: Speed-accuracy trade-offs in video classification. In *Proceedings of the European Conference on Computer Vision (ECCV)*.

Aoxiong Yin, Tianyun Zhong, Li Tang, Weike Jin, Tao Jin, and Zhou Zhao. 2023. Gloss attention for gloss-free sign language translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2551–2562.

Kayo Yin, Kenneth DeHaan, and Malihe Alikhani. 2021a. Signed coreference resolution. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 4950–4961, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Kayo Yin, Amit Moryossef, Julie Hochgesang, Yoav Goldberg, and Malihe Alikhani. 2021b. Including signed languages in natural language processing. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 7347–7360, Online. Association for Computational Linguistics.

Youngwoo Yoon, Bok Cha, Joo-Haeng Lee, Minsu Jang, Jaeyeon Lee, Jaehong Kim, and Geehyuk Lee. 2020. Speech gesture generation from the trimodal context of text, audio, and speaker identity. *ACM Transactions on Graphics (TOG)*, 39(6):1–16.

Mike Zhang and Antonio Toral. 2019. The effect of translationese in machine translation test sets. In *Proceedings of the Fourth Conference on Machine Translation (Volume 1: Research Papers)*, pages 73–81, Florence, Italy. Association for Computational Linguistics.

Benjia Zhou, Zhigang Chen, Albert Clapés, Jun Wan, Yanyan Liang, Sergio Escalera, Zhen Lei, and Du Zhang. 2023. Gloss-free sign language translation: Improving from visual-language pretraining. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 20871–20881.

Hao Zhou, Wengang Zhou, Weizhen Qi, Junfu Pu, and Houqiang Li. 2021. Improving sign language translation with monolingual data by sign back-translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1316–1325.