

Accessible Translation from Standard Arabic to Moroccan Sign Language: A Rule-Based System for Low-Resource Sign Languages

Abdelhadi Soudi¹, Kristof Van Laerhoven², Mohammed Hannani¹

¹Ecole Nationale Supérieure des Mines de Rabat, ²University of Siegen

asoudi@enim.ac.ma, kvl@eti.uni-siegen.de, mohamed_hannani@yahoo.com

Abstract

Access to information is a major challenge for deaf communities, especially in emergencies where communication is often delivered through spoken or written language that may not be fully accessible to sign language users, and in regions where sign language linguistic resources and technological tools are limited. This paper presents a system for translating Standard Arabic text into Moroccan Sign Language (MSL) using a rule-based architecture enhanced with Large Language Model (LLM)-based lexical disambiguation. The system uses Arabic linguistic analysis, contextual word sense disambiguation (WSD), rule-based grammatical transformations, and video-based sign generation to produce sign language output. The main contribution of the system is the introduction of an automated disambiguation mechanism that resolves polysemous Arabic words without requiring manual user intervention. The approach builds on previously developed linguistic resources, including a bilingual Arabic-MSL lexical database with approximately 3000 entries and data collected within the framework of a large-scale MSL corpus project involving deaf participants from the 12 regions of Morocco. Preliminary evaluation shows that contextual disambiguation improves translation performance when ambiguous Arabic words correspond to different MSL signs. The proposed system contributes to the development of accessibility technologies for under-resourced sign languages and supports improved access to information for deaf communities.

1 Introduction

Research in cognitive neuroscience shows that the way language is perceived, processed, and produced depends on the interaction between the human brain, sensory organs, and the surrounding social environment (Dehaene, 2009). Spoken languages are primarily auditory and linear, whereas

sign languages are visual-gestural, spatial, and multimodal. Rather than relying on a standardized written form, they encode grammatical information through hand configuration, movement, orientation, location, facial expression, gaze, and body posture (Brentari, 2010; Sandler and Lillo-Martin, 2006). These fundamental modality differences have important implications for both linguistic analysis and computational processing.

Deaf and hard-of-hearing individuals often face significant barriers to linguistic accessibility. Communication support services, such as sign language interpreters, are frequently scarce, costly, or unavailable, which restricts access to education, employment, and public services. In Morocco, these challenges are evident. According to the 2014 census conducted by Morocco's High Commission for Planning, approximately 3.5% of the population (1,182,681 individuals) experience some degree of hearing loss. Among them, 0.2% (56,745 individuals) have profound deafness, 1.0% (347,386 individuals) report severe hearing difficulties, and 2.3% (778,550 individuals) experience moderate hearing difficulties (El Ouazzani, 2015; Lkhoulf, 2017). Despite the size of this population, access to educational materials and communication technologies is extremely limited.

The educational situation of deaf individuals in Morocco reflects these accessibility gaps. Most textbooks and teaching materials are produced in Standard Arabic or French, with virtually no content available in MSL. As a result, deaf learners often rely on informal instruction provided by community associations, where teachers may lack specialized training in deaf education and advanced proficiency in MSL. The absence of institutional programs dedicated to training sign language interpreters further exacerbates this situation, leaving many deaf individuals without reliable linguistic mediation in essential contexts such as healthcare, legal services, and public information (Soudi and

Vinopol, 2019).

Another major challenge is the limited linguistic documentation of MSL itself. Like many sign languages worldwide, MSL is under-documented, with relatively few linguistic descriptions or digital resources available. This lack of documentation makes it difficult to develop educational materials, linguistic tools, and assistive technologies for the deaf community. Linguistic research on sign languages in general still lags behind research on spoken languages. Indeed, very few sign languages possess comprehensive reference grammars comparable to those available for spoken languages (Palfreyman et al., 2015).

From an information and communication technology perspective, these challenges are amplified by the modality differences between spoken and signed languages. Spoken languages can be processed through text or speech recognition systems, whereas sign languages require visual representations such as animations, avatars, or video recordings for both analysis and generation. Accordingly, computational approaches designed for spoken language processing cannot be transferred directly to sign languages without significant adaptation.

Recent work has highlighted the importance of developing linguistic resources and assistive technologies for under-resourced sign languages (Efthimiou et al., 2020; Johnston, 2010). However, most computational research in sign language processing has focused on well-resourced languages (Camgöz et al., 2018). MSL is largely absent from this research landscape.

To address this gap, this paper presents a rule-based translation system that converts Standard Arabic text into MSL. The system integrates linguistic analysis of Arabic input, AI-assisted lexical disambiguation, rule-based grammatical transformation, and video-based sign generation. The approach is specifically designed for low-resource sign languages, where large parallel corpora required by data-driven machine translation systems are not yet available.

The main contributions of this work include the integration of AI-assisted lexical disambiguation within a constrained translation pipeline, the introduction of a fingerspelling fallback mechanism for handling out-of-vocabulary items, and the use of linguistic resources collected through a nationwide Moroccan Sign Language corpus initiative involving deaf participants, which helps identify the grammatical structure of MSL and supports

the generation of translations consistent with its linguistic features.

2 Related Work

Research on sign language processing has expanded considerably in recent years, and includes tasks, such as sign recognition, translation, avatar generation, and multimodal linguistic analysis (Bragg et al., 2019). Early sign language translation systems were primarily rule-based, relying on explicit linguistic transformations to convert spoken or written language structures into sign language. For example, Stein et al. (2006) proposed a statistical translation system for German Sign Language enhanced with morpho-syntactic information, and Morrissey and Way (2005) investigated example-based translation between English and Sign Language of the Netherlands (NGT). These approaches rely on explicit grammatical rules, bilingual dictionaries, and structured linguistic representations to generate sign sequences.

More recently, data-driven approaches have been widely used in machine translation research. These include Statistical Machine Translation (SMT) and Neural Machine Translation (NMT). While SMT systems learn translation patterns from parallel corpora using statistical models, NMT architectures, mainly those based on the Transformer model, learn contextual representations through deep neural networks. However, both SMT and NMT approaches require large annotated datasets, which are rarely available for most sign languages. This limitation is particularly evident in sign language translation. The PHOENIX14-T corpus (Forster et al., 2012), which consists of approximately 8,000 weather-forecast sentences, was among the earliest datasets for sign language translation research. It enabled the development of deep learning approaches, including the Transformer-based models introduced by Camgöz et al. (2020).

Although these models achieved promising results on PHOENIX14-T, their performance drops significantly when applied to more diverse datasets, which makes it difficult to generalize neural models beyond restricted domains.

Recent research has attempted to address these limitations by integrating multimodal and multitask learning architectures. Models such as SLTUNet (Zhang et al., 2023) combine sign-to-gloss and gloss-to-text modules, whereas other approaches incorporate intermodal representations

that link text, gloss, and video data. Pre-trained language models and Large Language Models (LLMs) have also opened new possibilities for low-resource and sign language translation. Models such as BERT (Devlin et al., 2019), mBART (Liu et al., 2020), and GPT-3 (Brown et al., 2020) have demonstrated strong contextual representation, multilingual transfer, and few-shot learning capabilities, respectively. Nevertheless, these approaches are highly dependent on the availability of large annotated datasets and often fail to account for regional variation and the multimodal nature of sign languages.

For under-resourced sign languages, corpus development and linguistic documentation are essential prerequisites for natural language processing (NLP) applications (Johnston, 2010; Crasborn and Hanke, 2003). Several large-scale sign language corpora have been created. These include the BSL Corpus (Schembri et al., 2013), the Auslan Corpus (Johnston, 2008), and the DGS Corpus for German Sign Language (Konrad et al., 2020). These projects have provided linguistic resources and methodological frameworks for documenting sign languages and supporting computational research.

However, MSL is largely under-documented. Early lexical resources contained only a few hundred signs (Wismann and Walsh, 1987), while more recent initiatives have focused on developing educational resources and digital sign databases. The scarcity of large-scale annotated datasets limits the use of data-driven approaches for Arabic-MSL translation.

In low-resource contexts, alternative approaches such as Example-Based Machine Translation (EBMT) have been explored. EBMT constructs translations by analogy with previously translated segments, making it more interpretable than neural approaches and potentially more suitable for languages with limited training data. Prior work on Brazilian Sign Language generation illustrates the use of avatar-based systems for sign output (DeMartino et al., 2017). More broadly, work on low-resource machine translation has examined trade-offs between rule-based and data-driven approaches (Ramesh and Sankaranarayanan, 2018; Bayón and Sánchez-Gijón, 2019).

Despite these advances, the lack of extensive corpora is a central obstacle to building NLP applications for sign languages. Hurskainen and Tiedemann (2018) argued that the exclusive reliance on large-scale neural approaches risks excluding the

vast majority of the world’s languages, approximately 99.8%, from computational development.

In light of these constraints, this work adopts a rule-based machine translation (RBMT) approach for Arabic-MSL translation. While MSL is under-resourced, a small but growing set of linguistic resources described in prior work (Souidi and Vinopol, 2019) provides an initial foundation for this approach. More precisely, a bilingual lexical database of approximately 3000 Standard Arabic words and their corresponding MSL video signs supports the translation process. In addition, the Madrasati-Signs platform (Souidi and Van Laerhoven, 2025) facilitates the lexical documentation of MSL through a crowdsourcing approach that involves deaf associations and deaf signers across Morocco. The platform allows contributors to record and validate sign videos for Arabic lexical items in a database categorized into 21 thematic domains: family, colors, home, food, clothing, time, sports and hobbies, body, emotions, geography, transportation, education, nature and weather, animals, health, mathematics, business and careers, media, arts, technology, the alphabet, life sciences, and physics). The platform supports the gradual expansion and verification of the bilingual sign database used by the system. An important advantage of the RBMT architecture is its modularity and extensibility. New lexical entries and grammatical rules can be incorporated incrementally without retraining the entire system. This design also facilitates the integration of MSL regional variations which are very highly based on the so far collected data.

3 Arabic-MSL Translation System

This section describes the architecture of the proposed system for translating Standard Arabic text into Moroccan Sign Language (MSL). The translation process follows a modular pipeline that includes Arabic linguistic analysis, contextual lexical disambiguation, rule-based grammatical transformations, and sign video generation. The overall architecture of the system is illustrated in Figure 1.

The translation pipeline begins with preprocessing of the Arabic input sentence, including normalization and tokenization. The tokens are then processed by an Arabic linguistic analysis module that performs morphological analysis using AraMorph (Buckwalter, 2002) and syntactic dependency parsing and morphological disambiguation

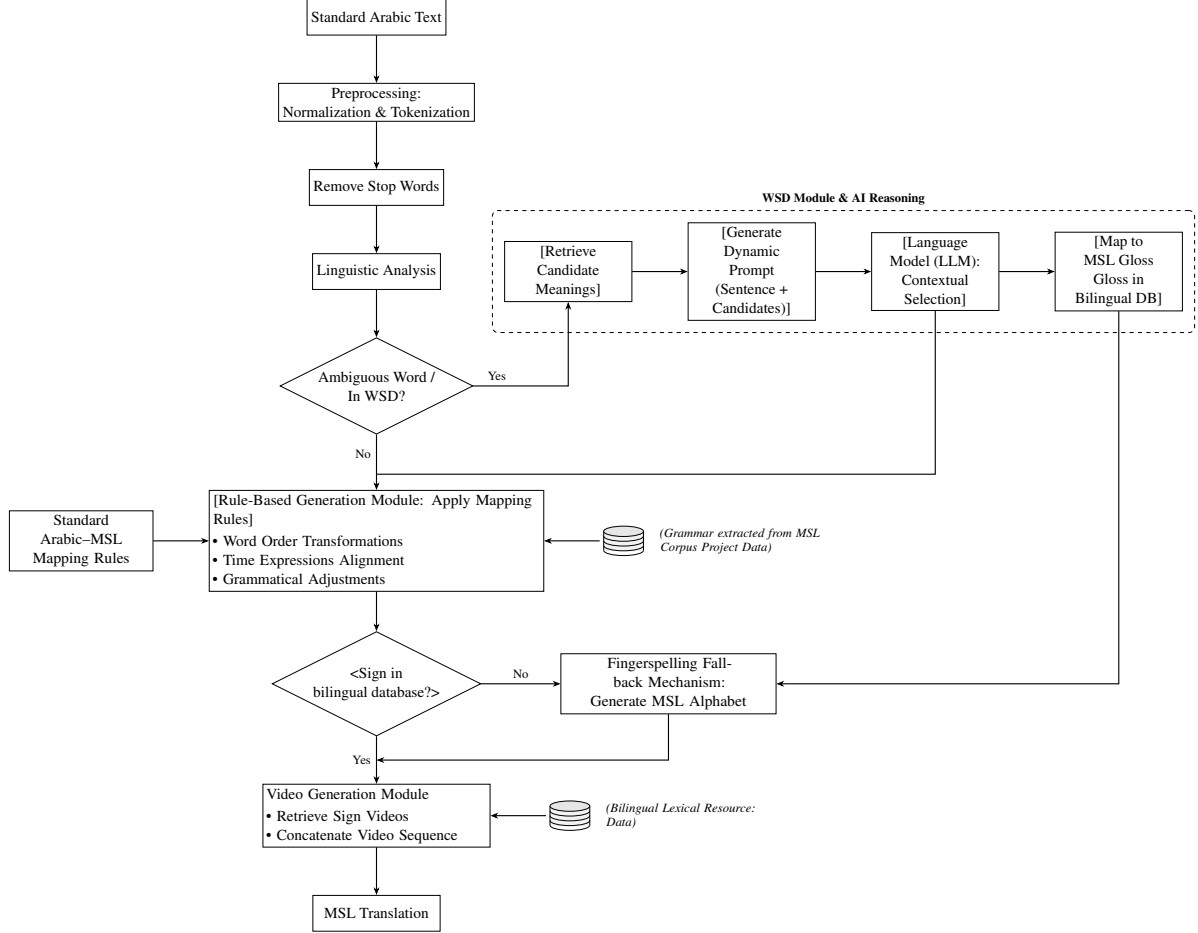


Figure 1: Architecture of the Standard Arabic–Moroccan Sign Language Translation System

using CamelParser (Shahrour et al., 2016). The resulting linguistic information is used to establish the morphological and syntactic context of the input sentence.

The system then checks each token against the WSD database in order to detect lexical ambiguity. Arabic contains many polysemous words whose interpretation depends strongly on context, and selecting an incorrect meaning can lead to an incorrect MSL sign.

In our previous work on Arabic–MSL translation, we handled lexical ambiguity through explicit user interaction (Souidi et al., 2025a). When the system encountered a polysemous word, it retrieved the list of possible meanings from a lexical resource and presented them to the user, who was required to manually select the correct interpretation before the translation process could continue. While this strategy ensured semantic correctness, it introduced a dependency on user intervention and prevented full automation of the translation pipeline.

To overcome this limitation, the present system introduces an AI-assisted lexical disambiguation

mechanism that automatically performs contextual interpretation. The system relies on the WSD dataset introduced by Kaddoura and Nassar (2024), which contains 100 frequently used polysemous Arabic words associated with 367 different senses. Each sense is accompanied by contextual sentence examples that illustrate the appropriate interpretation of the word in different semantic contexts. The dataset contains approximately 3,670 contextual sentence samples collected from domains, such as news, medicine, and finance.

When the system encounters a lexical item that exists in the WSD database, it retrieves the candidate meanings associated with that word and includes them in a prompt submitted to a language model along with the original sentence context. The model evaluates the semantic compatibility between the sentence and the candidate meanings and selects the most appropriate interpretation. The selected meaning is then mapped to the corresponding MSL gloss stored in the bilingual translation database. The model selects from predefined candidate meanings, which ensures that the chosen

interpretation corresponds to a sign representation available in the translation system. Once the lexical interpretation has been resolved, the system applies mapping rules and rule-based grammatical transformations that convert the Arabic sentence structure into the MSL grammatical structure. These rules account for differences between spoken and signed languages, which include, inter alia, variations in word order and the treatment of temporal expressions.

The translation system draws on two complementary linguistic resources: a Moroccan Sign Language (MSL) corpus and a bilingual Arabic–MSL lexical database that consists of 3000 Standard Arabic signs and their corresponding MSL video signs and graphics. The MSL corpus project is an ongoing initiative to document MSL through large-scale video data collection, carried out in collaboration with deaf associations across all twelve administrative regions of Morocco (Soudi et al., 2025b). Participants are selected based on sociolinguistic criteria, including age, gender, regional origin, and language acquisition background. The corpus initiative aims to collect recordings from 240 deaf signers distributed across the twelve regions of Morocco. During the pilot phase, 66 participants from four regions contributed to the initial data collection, producing approximately 27,91 hours of conversational and interview video recordings. These recordings are a valuable resource to understand MSL grammar and variation. The collected video data are annotated using the ELAN annotation environment, which allows time-aligned annotation of multimodal linguistic features such as glosses, hand configuration, facial expressions, and spatial structures (Wittenburg et al., 2006).

After the gloss sequence has been generated, the system produces the final sign language output through a video concatenation module. Each gloss corresponds to a sign video stored in the bilingual database. The system retrieves the relevant video signs and concatenates them according to the MSL sentence structure.

The disambiguation strategy can be illustrated by the Arabic word *عجين*, which may have different meanings depending on context. In a culinary context, such as *تحضير العجين للخبز* (“preparing the dough for bread”), the word refers to bread dough, which corresponds to an MSL sign. However, in a different context such as *يلعب الطفل بعجين الأطفال* (“the child plays with

modeling clay”), the word refers to children’s modeling clay (play dough). These meanings correspond to different MSL signs, as illustrated in Figure 2. When the system encounters the word *عجين*, it retrieves the candidate meanings from the WSD database and evaluates them within the sentence context. The AI-assisted disambiguation module then selects the interpretation that best fits the sentence before mapping it to the corresponding MSL sign representation.

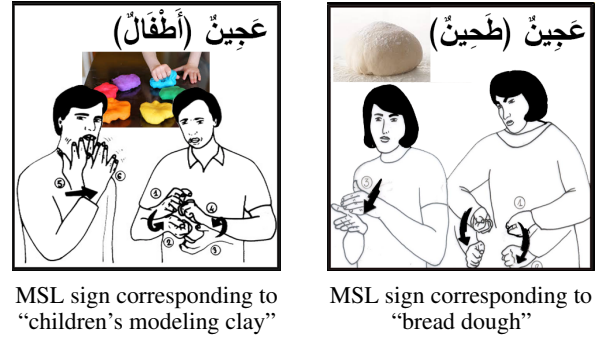


Figure 2: Lexical disambiguation example for the Arabic word “عجين” with two corresponding Moroccan Sign Language signs

To ensure that every word in the input sentence is translated, the system includes a fingerspelling fallback mechanism. If a lexical item does not exist in the Standard Arabic–MSL database, the system automatically fingerspells it.

4 Evaluation

Evaluating translation systems for low-resource sign languages is challenging due to the limited availability of standardized datasets and annotated evaluation benchmarks. Unlike well-resourced sign languages, MSL currently lacks large parallel corpora that could support large-scale quantitative evaluation. As a result, evaluation strategies must often rely on targeted test sets and qualitative evaluation of translation outputs.

To assess the effectiveness of the AI-assisted lexical disambiguation strategy, we analyzed the 3000-word Arabic–MSL bilingual database used by the system and identified 38 ambiguous Arabic lexical items. Among these ambiguous words, 15 are also represented in the WSD dataset used by the disambiguation module. Test sentences containing these ambiguous words were constructed to examine the system’s ability to select the appropriate meaning in different contextual environments.

The system processed these sentences through

the complete translation pipeline, including contextual lexical disambiguation and rule-based generation of Moroccan Sign Language gloss sequences. The results show that the AI-assisted disambiguation mechanism selects contextually appropriate meanings and maps them to the corresponding MSL signs. Compared with the earlier version of the system, which required manual user selection, the new approach enables the translation pipeline to operate fully automatically.

These observations indicate that the integration of AI-based contextual reasoning improves the translation process in cases where polysemous Arabic words correspond to different signs.

5 Limitations

A major limitation of the current system concerns the coverage of the available lexical and linguistic resources. Although the bilingual database contains approximately 3000 lexical entries, it does not yet cover the full vocabulary of Standard Arabic.

Similarly, the Word Sense Disambiguation dataset used in the system is not exhaustive and currently includes only a subset of ambiguous Arabic words. As a result, some cases of lexical ambiguity may remain unresolved when the corresponding lexical items are not present in the disambiguation resource.

Another limitation concerns the size of the MSL corpus. The amount of annotated data is still limited. Expanding the lexical database and enriching the annotated corpus will therefore be important steps for improving translation coverage and supporting future data-driven approaches for Arabic-MSL translation.

It is worth noting that the documentation of sign languages requires sustained long-term efforts in data collection, annotation, and validation. The approach adopted in this work recognizes the importance of involving the Deaf community from the beginning of the research process and supporting their leadership in the documentation of MSL data. This strategy aims to foster a sense of ownership of the linguistic resources within the Deaf community, which may support the long-term sustainability and continuous expansion of the MSL documentation effort.

6 Conclusion

This paper presented a rule-based Standard-Arabic MT system that is enhanced with AI-assisted lexical disambiguation. The system integrates Arabic linguistic analysis, contextual word sense disambiguation, rule-based grammatical transformations, and video-based sign generation to produce sign language output.

A major improvement over previous work is the introduction of an automated contextual disambiguation mechanism, which replaces the earlier manual ambiguity-resolution process and allows the translation pipeline to operate fully automatically.

The system also benefits from the results of the large-scale MSL corpus project, which contribute to a better understanding of MSL grammar, an important component of the MSL generation module that generates outputs that follow MSL linguistic features.

Future work will focus on expanding the lexical database and MSL annotation in order to further improve the quality and coverage of Arabic-MSL translation.

References

- María Do Campo Bayón and Pilar Sánchez-Gijón. 2019. [Evaluating machine translation in a low-resource language combination: Spanish-Galician](#). In *Proceedings of Machine Translation Summit XVII: Translator, Project and User Tracks*, pages 30–35, Dublin, Ireland.
- Danielle Bragg, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreaux, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, Christian Vogler, and Meredith Ringel Morris. 2019. [Sign language recognition, generation, and translation: An interdisciplinary perspective](#). In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19)*, pages 16–31. Association for Computing Machinery.
- Diane Brentari. 2010. *Sign Languages*. Cambridge University Press.
- Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33:1877–1901.
- Tim Buckwalter. 2002. Buckwalter arabic morphological analyzer version 1.0. Technical Report LDC2002L49, Linguistic Data Consortium.

- Necati Cihan Camgöz, Simon Hadfield, Oscar Koller, and Richard Bowden. 2018. [Neural sign language translation](#). In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7784–7793.
- Necati Cihan Camgöz, Oscar Koller, Simon Hadfield, and Richard Bowden. 2020. [Sign language transformers: Joint end-to-end sign language recognition and translation](#). In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10023–10033.
- Onno Crasborn and Thomas Hanke. 2003. Metadata for sign language corpora. In *Proceedings of the Corpus Linguistics 2003 Conference*, volume 16, pages 136–140. Lancaster University.
- Stanislas Dehaene. 2009. *Reading in the Brain: The New Science of How We Read*. Penguin, New York.
- José Mario DeMartino, Ivani Rodrigues Silva, Carmen Zink Bolognini, Paula Dornhofer Paro Costa, Kátia Mara Kumada, et al. 2017. [Signing avatars: making education more inclusive](#). *Universal Access in the Information Society*, 16(3):793–808.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. [BERT: Pre-training of deep bidirectional transformers for language understanding](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Eleni Efthimiou, Stavroula-Evita Fotinea, Thomas Hanke, Julie A. Hochgesang, Jette Kristoffersen, and Johanna Mesch, editors. 2020. *Proceedings of the LREC 2020 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*. European Language Resources Association (ELRA).
- Zakaria El Ouazzani. 2015. [The moroccan experience on disability statistics](#). Presentation at the 15th Annual Meeting of the Washington Group on Disability Statistics, Copenhagen, Denmark.
- Jens Forster, Christoph Schmidt, Thomas Hoyoux, Oscar Koller, Uwe Zelle, Justus Piater, and Hermann Ney. 2012. [RWTH-PHOENIX-Weather: A large vocabulary sign language recognition and translation corpus](#). In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC)*, pages 3785–3789, Istanbul, Turkey.
- Arvi Hurskainen and Jörg Tiedemann. 2018. [Rule-based machine translation from English to Finnish](#). In *Proceedings of the Second Conference on Machine Translation (WMT17)*, pages 310–319. Association for Computational Linguistics.
- Trevor Johnston. 2008. The Auslan corpus. In *Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora*, pages 96–101. European Language Resources Association (ELRA).
- Trevor Johnston. 2010. [From archive to corpus: The tagging of sign language corpora in ELAN](#). *Linguistics*, 48(5):1061–1100.
- Sanaa Kaddoura and Reem Nassar. 2024. [A comprehensive dataset for Arabic word sense disambiguation](#). *Data in Brief*, 54:110591.
- Reiner Konrad, Thomas Hanke, Gabriele Langer, Dolly Blanck, Julian Bleicken, Iлона 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](#). Universität Hamburg.
- Yinhan Liu, Jiatao Gu, Naman Goyal, Xian Li, Sergey Edunov, Marjan Ghazvininejad, Mike Lewis, and Luke Zettlemoyer. 2020. [Multilingual denoising pre-training for neural machine translation](#). *Transactions of the Association for Computational Linguistics*, 8:726–742.
- N. Lkhoulf. 2017. Disability statistics from the 2014 moroccan census. In *Regional Meeting on Disability Measurement and Statistics in support of the 2030 Agenda for Sustainable Development and the 2020 World Population and Housing Census Programme*, Muscat, Oman.
- Sara Morrissey and Andy Way. 2005. [An example-based approach to translating sign language](#). In *Proceedings of the Workshop on Example-Based Machine Translation (MT Summit X)*, pages 109–116, Phuket, Thailand.
- Nick Palfreyman, Keiko Sagara, and Ulrike Zeshan. 2015. Methods in carrying out language typological research. In Eleni Orfanidou, Bencie Woll, and Gary Morgan, editors, *Research Methods in Sign Language Studies: A Practical Guide*, pages 173–192. Wiley-Blackwell, Chichester, UK.
- Sree Harsha Ramesh and Krishna Prasad Sankaranarayanan. 2018. [Neural machine translation for low resource languages using bilingual lexicon induced from comparable corpora](#). In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop (NAACL-HLT SRW)*, pages 112–119.
- Wendy Sandler and Diane Lillo-Martin. 2006. *Sign Language and Linguistic Universals*. Cambridge University Press.
- Adam Schembri, Jordan Fenlon, Ramas Rentelis, Sally Stamou, and Kearsy Cormier. 2013. [Building the](#)

- British Sign Language Corpus. *Language Resources and Evaluation*, 47(4):1179–1207.
- Anas Shahrour, Salam Khalifa, Dima Taji, and Nizar Habash. 2016. CamelParser: A system for Arabic syntactic analysis and morphological disambiguation. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*, pages 228–232. The COLING 2016 Organizing Committee.
- Abdelhadi Soudi and Kristof Van Laerhoven. 2025. [Madrasati-signs: A hybrid crowdsourcing approach to creating a Moroccan Sign Language school dictionary for general and STEM terms](#). In *Proceedings of the 2025 IEEE 8th International Congress on Information Science and Technology (CiSt)*, pages 1–6. IEEE.
- Abdelhadi Soudi and Corinne Vinopol. 2019. [Educational challenges for deaf and hard-of-hearing children in Morocco](#). In Harry Knoors, Maria Brons, and Marc Marschark, editors, *Deaf Education Beyond the Western World: Context, Challenges, and Prospects*. Oxford Academic, New York.
- Abdelhadi Soudi, Corinne Vinopol, and Kristof Van Laerhoven. 2025a. [Bilingual resources for Moroccan Sign Language generation and Standard Arabic skills improvement of deaf children](#). In *Proceedings of the 18th Workshop on Building and Using Comparable Corpora (BUCC)*, pages 1–9, Abu Dhabi, UAE. Association for Computational Linguistics.
- Abdelhadi Soudi, Corinne Vinopol, and Kristof Van Laerhoven. 2025b. [Toward the creation of a large-scale Moroccan Sign Language corpus](#). *Sign Language Studies*, 25(4):741–767.
- Daniel Stein, Jan Bungeroth, and Hermann Ney. 2006. Morpho-syntax based statistical methods for sign language translation. In *Proceedings of the 11th Annual Conference of the European Association for Machine Translation (EAMT)*, pages 169–177, Oslo, Norway.
- Lynn Wismann and Michael Walsh. 1987. *Signs of Morocco*. Peace Corps Morocco.
- Peter Wittenburg, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 2006. ELAN: A professional framework for multimodality research. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, pages 1556–1559. European Language Resources Association (ELRA).
- Biao Zhang, Mathias Müller, and Rico Sennrich. 2023. [SLTUNET: A simple unified model for sign language translation](#). In *Proceedings of the Eleventh International Conference on Learning Representations (ICLR)*.