Norwegian Sign Language: Overview of Resources and Experiments with Automatic SignWriting Transcription

Elisabeth Othamar, Yves Scherrer
University of Oslo
Oslo, Norway
elisaot@math.uio.no, yves.scherrer@ifi.uio.no

Abstract

Norwegian Sign Language (NTS) remains an under-resourced sign language despite its official recognition in Norway since 2022. The limited availability of structured, reusable, and publicly accessible datasets continues to hinder both linguistic research and the development of sign language technologies such as recognition and translation systems. This paper presents an overview of existing datasets and potential data sources for NTS, categorizing them by accessibility, format, and suitability for computational research. We further discuss legal, ethical, and practical considerations related to data reuse, including copyright and privacy constraints. In addition, we report on a series of pilot experiments exploring alternative data acquisition strategies, including dictionary videos, SignWriting resources, and broadcast news material. These preliminary experiments explore whether automatic SignWriting transcription can serve as an intermediate representation for NTS, and examine its potential role in sign identification within continuous signing. The aim of this work is both to document ongoing efforts and to support future initiatives toward the sustainable development of NTS resources.

Keywords: Norwegian Sign language, sign language recognition, sign language translation, data scarcity

1. Introduction

Sign language resources are a critical foundation for linguistic research and the development of sign language technologies. Resources for sign languages are being collected (e.g. see entries in *The Sign Language Dataset Compendium* by Kopf et al. (2022)). However, many languages are still low-resourced. Norwegian Sign Language (*norsk tegnspråk*, NTS) is one such example, with only lexical entries in *The Sign Language Dataset Compendium*[1]. Although NTS was officially recognized as the national sign language of Norway in 2022 (Halvorsen, 2025), the availability of structured and reusable datasets for research purposes remains limited.

Systematic corpus work on Norwegian Sign Language (NTS) began in the mid-2010s, although earlier datasets were collected as part of individual research projects (e.g., in 2007) (Jantunen et al., 2024), but there are currently no publicly available datasets suitable for large-scale annotation or computational processing of NTS. Existing datasets are restricted to private use (Svendsen and Kadry, 2023; Uddin et al., 2025) and/or lack the necessary annotations. Recent work has explored the use of broadcast media as a data source (Høgset, 2024), but this introduces various challenges related to data quality, copyright, and reuse.

In response to these limitations, we explore an alternative route: leveraging partially structured resources, that were not originally designed for computational research, and leveraging existing language independent processing tools. These include single-sign dictionary videos, SignWriting lexicons, and publicly available news broadcasts in NTS. By combining these sources, we present initial experiments examining the feasibility of using automatic SignWriting transcription as an intermediate representation for NTS, and outline ongoing work on its potential application to sign identification in continuous signing. This paper therefore provides both an overview of available data sources and a pilot investigation into their practical usability for sign language technology development using language independent tools/frameworks, in the hope that such methods can be used for other low-resource sign languages.

2. Background and Related Work

Research on sign language resources has expanded significantly in recent years, reflecting both linguistic and technological developments. Importantly, current datasets differ along several key dimensions, including the type of signing captured, the form of annotation provided and the availability (e.g. see Núñez-Marcos et al. (2023) for publicly available resources and article[2] by Amit Moryossef and Yoav Goldberg, with resources of varying availability and type of signing).

[1] https://www.sign-lang.uni-hamburg.de/lr/compendium/language/nsl.html

[2] https://research.sign.mt

Proceedings of the 12th Workshop on the Representation and Processing of Sign Languages, pages 408–418
16 May 2026.

2.1. Aspects of Sign Language Resources

First, a central distinction concerns single-sign videos and continuous-signing videos. **Single-sign datasets** typically contain isolated lexical items produced in (or close to) citation form (i.e., one sign per video). Such datasets are commonly used to model sign lexemes and are often aligned with dictionary-style representations of sign vocabularies (Bragg et al., 2019; Koller, 2020). In contrast, **continuous-signing datasets** consist of fluent, multi-sign utterances forming sentences or longer discourses. Continuous signing introduces additional linguistic phenomena such as coarticulation, prosodic structuring, non-manual markers, and variation in signing speed, which make automatic transcription substantially more challenging than (isolated) single-sign data (Bragg et al., 2019; Koller, 2020; Núñez-Marcos et al., 2023). However, such datasets are essential for studying naturalistic language use and for developing NLP technologies such as continuous sign language recognition and translation (e.g., Duarte et al., 2021).

Second, sign languages can be represented using different annotation systems (Bragg et al., 2019). Most commonly, sign language data is annotated with **glosses**. Glosses are text-based representations of signs and often correspond to words of an oral language (Koller, 2020). For example, a phrase in the Australian Sign Language could be glossed as: *ASK-him STUDENT WHERE IX-he LIVE*, meaning *Ask the student where he lives* (Núñez-Marcos et al., 2023). While glosses are not full translations, they provide convenient symbolic labels for signs and are widely used in linguistic corpora and computational pipelines. However, glossing abstracts away much of the phonological and spatial structure of signs, including handshape, movement, location, and non-manual features. In addition, glosses are language- and dataset-specific: sign–gloss mappings valid in one sign language may not hold for another, or may vary across different datasets for the same language (e.g., Bragg et al., 2019; Koller, 2020). In contrast, **SignWriting** (Sutton, 1990)[3] and **HamNoSys** (Hanke, 2004)[4] are writing systems specifically designed for sign languages. These systems visually encode handshape, movement, orientation, location, and non-manual features in a two-dimensional layout with language independent symbols (see Table 5 for a SignWriting example). Compared to glossing, they offer a more fine-grained and language-internal representation, and can be applied to any Sign Language. Lastly, it is also possible to represent signs by natural language **descriptions** of both their manual and non-manual components. Such descriptions, typically in English, can be easily generated automatically with video LLMs (see e.g., Asasi et al., 2025). They are also largely language-independent and can therefore serve as an intermediate representation in NLP tasks.

Third, a component of many sign languages is **fingerspelling**. Fingerspelling involves the manual representation of written-language letters using specific handshapes, and is commonly used for proper names, technical terms, borrowings, or lexicalized items. From a computational perspective, fingerspelling poses unique challenges due to rapid articulation, coarticulation effects, and variation between fully spelled and lexicalized forms. Including fingerspelling in sign language research is crucial for building accessibility tools and applications that can accurately process names, specialized vocabulary, and other elements not covered by standard sign notation (Georg et al., 2025; Tanzer, 2025).

Across sign languages, the development of annotated large-scale datasets, particularly those containing continuous signing, has enabled progress in both linguistic research and NLP technologies (Power et al., 2025; Cihan Camgöz et al., 2020). These advances include improved automatic sign recognition, alignment between video and representations, and sign language translation systems (Koller, 2020; Núñez-Marcos et al., 2023). As resources continue to diversify in terms of modality and annotation scheme, careful consideration of these representational distinctions remains crucial for both corpus design and computational modeling.

2.2. Sign Language Research in Norway

NTS has received limited attention in terms of dataset creation. At the Western Norway University of Applied Sciences (HVL) in Bergen there is a research group focused on Sign Language, linguistics and interpreting.[5] Also, at the University of Bergen, there is a research group, TEIKN, researching both linguistics and the use of technology in processing Sign Languages.[6] Furthermore, at the Norwegian University of Science and Technology (NTNU) in Trondheim, there is a research project investigating how Deaf signers use NTS across linguistic and social contexts, while develop-

[3]https://www.signwriting.org/about/what/what02.html

[4]https://www.sign-lang.uni-hamburg.de/dgs-korpus/hamnosys-97.html

[5]https://www.hvl.no/en/research/group/TOLK/

[6]https://www4.uib.no/en/research/research-groups/teikn

ing a corpus and lexical database.[7] Researchers at the Department of Applied Data Science at Noroff University College in Kristiansand, investigated image classification models for NTS recognition (Svendsen and Kadry, 2023). Finally, SINTEF, an independent Norwegian research institute, is involved in a project aiming to develop an app for Sign Language translation for NTS.[8]

So far, existing efforts have primarily focused on linguistic documentation and research or small-scale educational resources, such as a publicly available sign dictionary, leaving a gap between available materials and the needs of computational research. Thus, the need for more NLP research and dataset creation for NTS is evident.

3. Norwegian Sign Language Datasets

This section reviews relevant resources for Norwegian Sign Language, starting with datasets that were explicitly created for NTS research or documentation, and continuing with alternative resources that primarily target the Deaf community but can be leveraged for computational experiments.

3.1. Curated Datasets

The *Norwegian Sign Language Corpus* encompasses four different datasets that were collected at NTNU between 2007 and 2024. These corpora were created for linguistic research and are in the process of being annotated. The resulting ELAN (EAF) files provide multi-tiered annotations for each signer, including right- and left-hand glosses, non-manual markers (body, face, head, gaze, mouthing), translations into Norwegian and English, and interactional features, allowing for detailed multimodal linguistic analysis. These datasets differ in the availability of ELAN files and their licensing:

- **Halvorsen (2012)**[9] – Collected in 2007, contains four deaf signers, two males and two females with varying age, which all have other deaf family members. The signers were asked to retell one children's book and answer one question. The dataset consists of 8 videos with a total duration of 18 minutes. The dataset is released under the CC BY-NC-SA 4.0 license. There are no annotations or translations available for this dataset. Additionally, all the signers lived in the central Eastern Norway and the videos may only capture one of the different dialects of NTS.

- **Pilot Corpus (Conversations)**[10] – The projectstarted in 2015 and aimed to document the language practices of elderly deaf signers and to introduce corpus research to the community. Seven elderly deaf signers from Oslo, Bergen, and Trondheim were recorded over two days at the OsloMet University. Participants engaged in dyadic and group conversations with other signers and deaf project team members. The total duration of this dataset is 5.5 hours. Parts of the dataset are available for non-commercial research with attribution, subject to approval of a research plan, compliance with privacy requirements, and restrictions against redistribution or creating derivative works. To the authors' knowledge, no translations are publicly available, though ELAN files exist for private recordings from this project (and for some Depicting Perspective videos) used in Ferrara et al. (2023).[11]

- **Depicting Perspective**[12] – Collected from 2017 to 2018, this dataset is a follow-up to previous L2 (sign language acquired as a second language) signing research. It captures how signers establish and maintain visual perspectives. Data were collected in Oslo, Bergen, and Trondheim with 21 deaf participants engaged in 13 naturalistic conversations (30–40 minutes each) with deaf or hearing native signers as interlocutors, totaling 7.5 hours of video. This dataset is licensed in the same manner as the Pilot Corpus. Some videos have additional restrictions, and their availability may be limited. Private recordings from this project and corresponding ELAN files[13] were also used in Ferrara et al. (2023).

- **Retellings and public events** – Collected from 2019 to 2024 as part of the **Language Ecology** project[14]. It contains retellings of three selected children's books and public presentations, recorded with 87 Norwegian Sign Language signers across multiple cities to study conversational interaction, multilingualism, and iconicity. The video duration is unknown. Retellings and Public Events is allow-

[7] https://www.ntnu.edu/isl/nts-language-ecology

[8] An interview about the SINTEF project can be found here.

[9] http://hdl.handle.net/11509/141

[10] http://hdl.handle.net/11509/147

[11] Some example ELAN files can be accessed via the OSF overview: https://osf.io/q8vf3/overview. Access to the full ELAN annotations may require contacting the corpus manager.

[12] http://hdl.handle.net/11509/144

[13] Some example ELAN files can be accessed via the OSF overview: https://osf.io/m4gh8/overview.

[14] http://hdl.handle.net/11509/155

Dataset	Format	Size	Duration	Availability
Halvorsen	Video	118 MB	18 minutes	Non-commercial
Pilot Corpus	Video	8,74 GB	5,5 hours	Upon approval
Depicting Perspective	Video	9,8 GB	7,5 hours	Upon approval
Retellings and public events	Video	13,8 GB	Unknown	Non-commercial
Conversations and private meetings/events	Video	68 GB	Unknown	Upon approval
Alphabet Dataset	Images	418 MB (24.300 images)	–	Available
Numbers Dataset	Video	Unknown	Unknown	Private

Table 1: Overview of known Norwegian Sign Language datasets.

ing non-commercial use, adaptation, and sharing with attribution, while requiring derivative works to be released under the same license.

- **Conversations and private meetings/events** – Is also part of the **Language Ecology** project, and contains conversations, discussions of deaf-related issues, private meetings, and events recorded with 87 Norwegian Sign Language signers (aged 22–91) across multiple cities. Sessions, typically 2.5 hours in pairs, capture naturalistic conversational interaction to study multilingualism, reference, and iconicity. The total duration of these videos is unclear. Conversations and Private Meetings/Events is distributed under a CLARIN RES license, allowing non-commercial research use with attribution, requiring submission of a research plan and approval by the corpus manager, and prohibiting redistribution or sharing of derivatives. Some data have additional restrictions and must be requested separately.

In sum, the varying licenses and availability of ELAN files makes this corpus unsuitable for data-demanding processes such as sign language translation and recognition. In the context of the Norwegian Sign Language Corpus, a sign bank for NTS was also collected, but it is not publicly available.[15]

Two additional datasets have been reported in the recent literature. One consists of still images of hands representing fingerspelling of the Norwegian **alphabet** (Svendsen and Kadry, 2023), while another one (part of the SINTEF initiative) contains short video recordings of the **numbers** 0–10 in NTS (Uddin et al., 2025). These datasets are limited in size and scope. The data for the numbers is private, however, the article for alphabet recognition indicate the data is obtainable, and this webpage contains the data under the CC BY 4.0 licence. The characteristics of these datasets are summarized in Table 1.

[15] https://signbank.cls.ru.nl/datasets/NTS/

3.2. Alternative Data Sources

Beyond curated datasets, several digital resources intended for the Deaf community contain valuable NTS material. These include broadcast news, educational videos, and online sign language platforms. While not designed for research, such sources may be leveraged to construct datasets under appropriate legal and ethical frameworks.

One such resource for creation of a single-sign dataset is the online NTS **dictionary**.[16] This dictionary is part of a resource called *Tegnbanken*[17], which states that the material is available under the CC BY-NC-ND 4.0 license.

SignPuddle[18] is a community-driven initiative to make SignWriting accessible to Deaf communities worldwide. SignPuddle hosts language-specific SQLite databases containing SignWriting entries paired with glosses or written-language translations. The Norwegian SignPuddle database contains over 6000 entries.[19] Such databases can serve as valuable parallel data for experiments involving SignWriting transcription, gloss prediction, or cross-modal mapping. However, because they are often collaboratively developed and independently maintained, questions related to quality, documentation, version control, and long-term sustainability must be carefully considered when incorporating them into reproducible NLP research pipelines.

A recent master's thesis explored the construction of a continuous-signing dataset based on **news broadcasts** aligned with corresponding subtitles (Høgset, 2024), taken from programs of the national broadcaster of Norway, NRK. While this approach increases coverage, it also introduces challenges related to alignment accuracy, bias, and copyright

[16] https://www.minetegn.no/Tegnordbok-2016/tegnordbok.php

[17] https://www.minetegn.no/Tegnbanken-2016/index.php

[18] Info on the people behind SignPuddle can be found at https://steveslevinski.me/#infrastructure.

[19] https://signpuddle.com/client/#!/dictionary/nsl-NO-dictionary-public

restrictions.[20]

Besides news broadcast, there exists a large online sign language platform called *tegn.tv*[21], whose objective is to aggregate and broadcast sign language content produced by various contributors. Many of the videos are created with accessibility in mind and include spoken Norwegian and/or subtitles; however, the platform does not explicitly guarantee that all videos provide subtitles, and coverage may therefore vary across content. The multitude of contributors makes it difficult to determine the copyright situations and conditions of reuse. However, in contrast to the videos at NRK, this content is largely made by the Deaf for the Deaf community, leading to a more representative language content in the videos.

In sum, broadcast media offers continuous signing and a broad vocabulary but is typically subject to strict copyright restrictions, whereas other sources may yield good quality single sign content, with the risk of differing from the same sign in the "wild" or simply being the effort of individuals not fluent/educated in sign language. These tradeoffs must be carefully considered when using such sources for research.

4. Going the Alternate Route

We present some preliminary experiments that rely on three alternative data sources: single-sign videos from the Tegnbanken online dictionary, the SignPuddle database for SignWriting, and a set of continuous-sign NRK news broadcasts. The goal of these experiments is to assess whether language independent annotation frameworks (such as SignWriting) and pretrained language independent annotation systems (such as those found in the Sign Language Processing open-source tools[22] by Amit Moryossef and Yoav Goldberg) can be leveraged for NTS in the absence of large manually annotated training data.

We evaluate how well automatic SignWriting transcription tools work for NTS. Prospectively, we aim to investigate if transcription to SignWriting (or to other language-independent sign representations) is useful for identifying single signs within continuous-sign videos, utilizing the SignWriting simply as a mapping to the dictionary gloss/translation.

4.1. Data Collection

The Tegnbanken dictionary is owned by Statped[23], and to obtain the single-sign videos, the authors programmatically harvested the website. An XML endpoint was queried to retrieve structured metadata, from which all video filename identifiers were extracted. For each identifier, a canonical URL pointing to the corresponding MP4 file was deterministically constructed. Videos were downloaded via parallelized HTTP requests with streaming transfer, file-integrity checks, and automatic handling of partial or failed downloads. Previously acquired files were skipped to ensure reproducibility and efficiency. The resulting dataset consists of 9119 locally stored lexical sign videos suitable for downstream tasks such as sign language recognition. The code can be found at https://github.com/elisaot/Code_NTS.git.

The NTS SignWriting reference is available as an individual SQLite database in the SignPuddle.[24] The database contained 6362 entries with SignWriting and corresponding gloss/translation to written Norwegian.

For the acquisition of NRK's news broadcasts it was decided to download one year of news, to maximize the vocabulary. We constructed the sign language news corpus by programmatically downloading episodes of NRK *Tegnspråknytt* from the NRK online media platform using the yt-dlp framework.[25] For each month URL, playlist metadata were extracted and automatically inspected for the availability of Norwegian subtitles. Only videos containing Norwegian subtitles were retained to ensure aligned written-language text. Selected episodes were downloaded in MP4 format together with their subtitle files, while previously acquired files were skipped to maintain reproducibility and efficiency. A structured download log was generated documenting episode identifiers and subtitle types. Due to copyright restrictions, we do not share the downloading code nor the videos themselves, but we release the processing log and the URLs used at https://github.com/elisaot/Code_NTS.git. The resulting dataset consists of 246 news broadcasts, with an average length of 4.2 minutes. Lastly, we automatically segmented the video material into subtitle-aligned clips by parsing time-stamped subtitle files and extracting corresponding video intervals. Subtitle times-

[20]NRK indicates that their products may be used for educational and research purposes: https://info.nrk.no/tv-spons-og-salg-av-nrks-innhold/arkiv/#innhold-til-undervisning-eller-forskning

[21]https://tegn.tv

[22]https://github.com/sign-language-processing

[23]Statped is the Norwegian national service for special education, providing support, resources, and expertise for individuals, including sign language users. https://www.statped.no

[24]The database for NTS (sgn69.db) can be found at http://signbank.org/swserver_data/puddle/.

[25]https://github.com/yt-dlp/yt-dlp?tab=readme-ov-file

Dataset	Format	File size	Amount	Availability
Tegnbanken Dictionary	Video + Gloss	2.72 GB	5.29 hours	Non-commercial
SignPuddle database	Gloss + SignWriting	1.8 MB	6362 entries	Open source (MIT)
NRK News Broadcasts	Video + Text	38.92 GB	17.35 hours	Copyrighted*

Table 2: Overview of the created/collected Norwegian Sign Language datasets.(* According to the NRK website, content for documentation and research purposes can be ordered via the National Library of Norway or through a local library.)

tamps were processed using pysrt, and video segments were extracted with moviepy. While this method enables efficient alignment of signed content with written-language subtitles, it relies on subtitle timing and does not guarantee precise linguistic segmentation. Additionally, the news broadcasts are interpreted from written Norwegian text and the person interpreting may be an L2 signer of NTS, introducing biases.

Table 2 shows the amount of data acquired from the three sources.

4.2. Inferring a Reference Dictionary With SignWriting Annotations

In order to evaluate automatic video-to-SignWriting tools, we need a reference dataset that associates single-sign videos with their SignWriting representations. The videos contains detailed gloss annotations (ID glosses) that distinguish between multiple sign variants (e.g., GLOSS-1, GLOSS-2). To enable alignment with the SignWriting data, these glosses are normalized into a simplified gloss set, where variant distinctions are collapsed (e.g., GLOSS-1, GLOSS-2 → GLOSS). The resulting simplified glosses serve as a pivot representation for cross-modal alignment between video and SignWriting.

As in written and spoken languages, multiple distinct signs in a sign language may share the same meaning, which is seen in our single-sign videos. Subtracting away their distinctions for alignment, performing no manual cleaning of these matches due to lack of SignWriting expertize, led to substantial lexical overlap, and consequently substantial multiplicity, in our reference dictionary.

Analysis of this revealed; In the SignPuddle database for NTS, 6362 entries were available, but only 1949 had a unique gloss. Of the 9119 single-sign videos examined, only 5598 of the pivot glosses occurred uniquely. When cross-referencing the two resources, only 2316 single-sign videos had a corresponding entry in SignPuddle. Thus, as a result of multiplicities in both the videos and the SignPuddle, the resulting dataset contained 5694 rows with only 519 unique glosses. For example, the gloss VANN ('water') showed the highest level of multiplicity across the datasets, yielding 66 entries in the reference dictionary.

Owing to limited expertise in both NTS[26] and SignWriting, manual verification to eliminate potential false matches between signs and SignWriting representations was not feasible. Consequently, while the dataset provides a valuable starting point for pose-to-SignWriting or gloss-based experiments, the presence of duplicate and potentially misaligned entries must be considered. For our project, this implies that any evaluation of automatic video-to-SignWriting tools using this dictionary should be regarded as preliminary and indicative rather than definitive, reflecting the quality of the underlying alignments rather than solely the performance of the models.

4.3. Automatic SignWriting Transcription

SignWriting transcription proceeds in three steps: first, the videos are transformed into **poses**, then the poses are segmented and lastly, the poses are transcribed into sequences of SignWriting signs.

For the first step, we make use of the Sign Language Processing tool Pose[27] by Amit Moryossef, Mathias Müller, and Rebecka Fahrni. An open-source toolkit that defines a standardized .pose format for storing temporally aligned body, hand, and facial keypoints extracted from video. The library provides utilities for reading, writing, preprocessing, and visualizing pose sequences, facilitating reproducible data preparation and consistent pose-based modeling in sign language recognition and translation tasks.

We applied the Linguistically Motivated Sign Language Segmentation (Moryossef et al., 2023) pipeline (utilizing the E1s model) to the .pose files. While not strictly necessary for single-sign videos, we applied the same tool to both continuous-sign and single-sign poses to maintain consistent processing and annotation.

The SignWriting transcription library[28], by Amit

[26]One of the authors is currently in the early stages of acquiring NTS as a second language (L2 signer).

[27]https://github.com/sign-language-processing/pose

[28]https://github.com/sign-language-processing/signwriting-transcription/tree/main

Moryossef, Rotem Zilberman, and Ohad Langer, provides a language-independent pre-trained model for pose-to-SignWriting transcription. The library defines a command-line interface that takes .pose files and corresponding ELAN files, normalizes and aligns the pose input, and runs a pre-trained neural transcription model to predict SignWriting symbol sequences.

Since the dictionary videos refer to single signs, we expect the transcription to mostly produce one SignWriting sign per video. However, we noted early on that dictionary videos were annotated with several SignWriting sequences, i.e. on average a dictionary video was annotated as two consecutive signs. This can be due to poor segmentation, the slower and larger signing, compared to natural conversation, due to the educational purpose of the videos, or a combination thereof. Additionally, some signs are multisyllabic, giving the appearance of two signs. In an attempt to counter the slow and large signing, we experimented with a pre-processing pipeline that generates speed- and/or trim-modified versions of the videos matched with entries from the SignPuddle. The transformations, including a 25% increase in playback speed and removal of 0.5 seconds from both the start and end of each video, were applied using moviepy.[29] This approach facilitates data augmentation and efficient preparation of large video datasets, while (optionally) preserving the original files.

4.4. Results and Analysis

For the evaluation of the SignWriting transcriptions, we relied on the work from Moryossef et al. (2024) on automatic evaluation metrics for SignWriting, relying on the data extracted from SignPuddle as reference transcriptions.

For manual evaluation SignWriting was visualized using the *signwriting_to_clip_image*[30] function from the SignWriting evaluation framework of Moryossef et al. (2024). To interpret the individual symbols, we consulted both the online Sutton SignWriting character reference[31] and the corresponding dictionary videos.

The metric results, found in 3, are very similar to the any-to-any comparison reported in Moryossef et al. (2024), where all sign pairs (of 1.000 randomly selected signs) are compared regardless of alignment. This suggests that the automatic annotations do not reliably align with the reference SignWriting: the metrics for supposed matches are comparable to the baseline of unrelated signs. The stability of these results across the experimental conditions (original, speed, cut, both) indicates that modification do not substantially affect this outcome. An important consideration regarding the metric scores; the misalignment of SignWriting in our results can be caused/exaggerated by the possible mismatches in our reference dictionary.

We conducted a manual evaluation of the similarity between predicted annotations for the dictionary videos and references for the alphabet, digits 0–10, and three randomly selected examples. We observed that mismatches between the video and the reference were particularly common in the alphabet. This is partly due to NTS including both one-handed and two-handed alphabet variants (Ferrara et al., 2023), which were not distinguished in our simplified gloss matching. Additionally, individual letters also had multiple similar variants, especially in the SignWriting references. Thus, in total, we had a manual evaluation dataset of 347 matches, where approximately half were either mismatches between videos and reference, videos without similar reference, or references without similar video.[32]

The following evaluation of predicted annotation quality considered videos and references that matched visually. For the alphabet, each of the four approaches (three modifications and the original) outperformed the others in approximately one-quarter of the cases. For digits 0–10, some handshape predictions were correct by chance, as the model tended to default to the most generic handshape (e.g., “5”). Finally, in the randomly selected examples, we observed handshapes and movements that were closer to the references, primarily in the original video predictions. However, there

[29] https://github.com/Zulko/moviepy

[30] https://github.com/sign-language-processing/signwriting-evaluation/blob/main/signwriting_evaluation/metrics/clip.py

[31] Which can be found here

[32] This manual verification was performed by one of the authors, who is currently acquiring NTS as a second language (L2 signer).

Metric	Original	Speed	Cut	Both
CHRF	0.2428	0.2427	0.2434	0.2431
CLIPScore	0.8708	0.8712	0.8694	0.8701
SymbolsDistances	0.3269	0.3265	0.3188	0.3197
TokenizedBLEU	0.0845	0.0857	0.0841	0.0843

Table 3: Mean metrics for the different experimental conditions on single-sign videos.

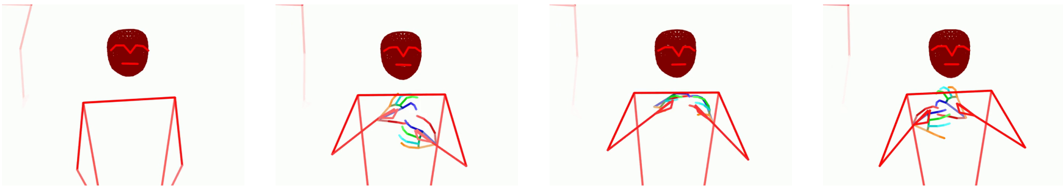


Table 4: Frames from the pose file for the NTS sign to gloss LENKE(girlander) (’chain(garlands)’).

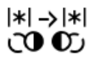
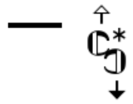
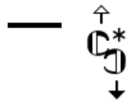
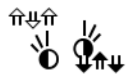
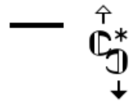
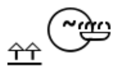
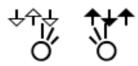
Reference	Original	Speed	Cut	Both
				
				

Table 5: SignWriting reference and predicted annotations for the NTS sign with ID-gloss LENKE(girlander) (’chain(garlands)’). From left to right: the reference SignWriting from SignPuddle; two consecutive predicted annotations for the original video; two consecutive predicted annotations for the video with increased playback speed; a single predicted annotation for the video trimmed by 0.5 seconds at the start and end; and two consecutive predicted annotations for the video that was both sped up and trimmed. (The horizontal line in the first predictions for original, speed and both modifications is the shoulders.)

are also inclusions of non-manual features that are not in the references from all video types.

Table 5 includes reference and predicted SignWriting for the NTS sign corresponding to gloss LENKE(girlander) (’chain(garlands)’), and table 4 shows the pose visualization of the sign. The automatic transcription produces multiple symbol sequences for a single sign, reflecting variations in signing speed, pose estimation noise, or poor segmentation. However, in this example, the sign is multi-syllabic and includes a hand-dominance switch; therefore, two separate predictions are reasonable.[33] The following is an evaluation of the different SignWritings (shown in Table 5) for the NTS sign corresponding to LENKE(girlander) (’chain(garlands)’):

- Reference (SignPuddle): The canonical SignWriting sequence captures the handshape, movement path, orientation, and spatial location consistently.
- Original video prediction (pose-to-SignWriting transcription): Handshapes are partially correct, almost forming a circle: the index and thumb should touch. The hands are forming the correct configuration in the first prediction. However, hand orientation is incorrect in both predictions, and the wrist rotation present in the sign is missing. One of the predictions introduces a non-manual feature (mouth movement) that is not present in the reference.
- Sped-up video predictions: The first prediction is identical to the first prediction of the original video. There is a slight improvement in that from the first to the second prediction the rotation of one of the hands seems to have been captured, but the second hand is missing, and the overall movement and placement are incorrect. A non-manual feature is again included that does not appear in the reference.
- Trimmed video predictions: Handshape is partially correct: the index and thumb form a circle, but the remaining fingers do not, as in the reference. Orientation appears correct, and the movement is predicted as up and down, whereas the actual sign involves wrist rotation and sideways movement. Additionally, the hands fail to interlock their circles.
- Sped-up and trimmed video predictions: The first prediction is identical to the original video’s first prediction. The second prediction resembles the trimmed video prediction, but the movement is now forward and backward rather than the reference. The hands still do not interlock correctly.

Overall, performance remained limited. Moreover, the experiments did not improve segmentation quality, as all video types produced an average of two predicted annotations per sign. This highlights the need for more robust segmentation and transcription methods, as well as careful consideration of video quality and signer variability, to produce

[33] We thank the reviewer for this observation.

accurate SignWriting representations for NTS.

5. Discussion and Open Challenges

Regarding our collected data we encounter several common challenges regarding Sign Language Processing; Our video sources and alternative sources are primarily protected by **copyright**, limiting research. Our reliance on non-research-oriented sources, such as broadcast material, introduces variability in recording conditions, domain bias, and **legal complexity**. Source selection also introduces an additional layer of bias.

Our interpreted news broadcasts involve **translation** from written (or spoken Norwegian) into NTS, which may influence linguistic structure and register. Furthermore, the objectivity norms of journalism may limit the range of facial expressions, affective markers, and body language represented in the data, making such material less representative of everyday signing "in the wild". Lastly, the interpreter may introduce additional bias depending on L1 or L2 signer status.

The dictionary-based resources from Tegnbanken have limited **representativeness**, as they present isolated signs in controlled conditions. However, unlike typical single-sign datasets created primarily for research purposes (Svendsen and Kadry, 2023; Uddin et al., 2025), Tegnbanken is designed for both Deaf users and learners of NTS. It functions as a dictionary, providing standard video representations and ID-glosses, making it particularly useful for language learning and teaching. For some concepts (simplified glosses), Tegnbanken provides multiple video variants, reflecting natural variation in NTS and the kinds of signs learners encounter in authentic communication. While this variability can complicate direct comparisons and merging with other sources, it also enables evaluation of automatic SignWriting annotations under conditions closer to real-world use, even if it results in lower or more variable evaluation metrics.

Similarly, the community-driven platform SignPuddle rely on voluntary contributions, which may vary in linguistic accuracy, completeness, and consistency. This introduces uncertainty and potential bias depending on the contributors' expertise. Thus, our method of making a reference dictionary from the two sources may perpetuate the limitations of SignPuddle. Additionally, our combined data lacks manual cleaning, resulting in mismatches and multiplicities. However, if the reference dictionary is later revised by experts in NTS and SignWriting, and the glosses are rather the more linguistically rich, fine-grained ID glosses of the videos, this dataset could be a valuable asset to NTS in the single sign setting.

6. Conclusion and Future Directions

This paper has presented an overview of existing datasets and potential data sources for Norwegian Sign Language (NTS), highlighting key limitations related to quality, licensing, privacy, and reuse. Moreover, the paper reports on a pilot experiment to leverage a language-independent SignWriting transcription model to annotate single-sign NTS videos. Our analysis shows that the automatically produced transcriptions rarely correspond to the ground truth and that the model is sensitive to minor variations in speed and length. However, it may still be useful for identifying keywords in the news broadcast videos. To evaluate this, we plan to annotate the continuous-sign videos with SignWriting transcriptions in the same way and investigate (a) how many signs can be retrieved from the automatically annotated dictionary, and (b) whether the glosses referring to the dictionary entries correspond to the words in the subtitles. These experiments are currently in progress. Further work will also explore alternative approaches such as the integration of video LLMs to produce language independent representations of NTS signs, to be used as mappings between video and gloss.

Our survey and pilot experiments further indicate that legal constraints, limited public availability of continuous signing data, and challenges related to documentation and anonymization pose substantial obstacles to methodological development in NTS recognition and translation research. Future work should thus prioritize collaborative, community-informed approaches to resource development, ensuring that dataset creation and publication, annotation practices, and evaluation methods are not only aligned with the linguistic expertise and perspectives of the Deaf community, but also enable research and development of NLP tools for sign languages. Such cooperation will be crucial for building robust, representative, and ethically grounded resources that support long-term progress in NTS research.

Lastly, further works for NTS will need to address the re-identification concerns regarding signing videos, which inherently contain **personally identifiable information** ((Xia et al., 2024; Bigand et al., 2021)). This risk is amplified in the small signing community of NTS, where the limited number of fluent signers increases the likelihood of recognition. Good solutions of this challenge may also improve the community's outlook, and involvement on dataset and technology creation (Bragg et al., 2020).

7. Limitations

Our pilot project, in addition to previously mentioned limitations and challenges, has several inherent limitations; Firstly, the pose transformation of our video data introduce its own set of limitations and bias from its reliance on MediaPipe Holistic[34] for keypoint extraction. Similarly, the segmentation model may introduce limitations and biases stemming from both the model itself and its training data. In addition, the segmentation model does not incorporate facial keypoints as inputs. Finally, the SignWriting annotation model may carry similar limitations, resulting in multiple stages of the pipeline being affected by methodological constraints and underexplored biases.

Our future use of the automatically annotated SignWriting of single signs is for identifying keywords in news broadcast videos. With this approach, we introduce an additional limitation: we consider intermediate representations (SignWriting) acceptable if they are predicted consistently. However, consistent predictions do not necessarily imply linguistic correctness, as our results in Section 4.4 demonstrate. Due to our limited expertise in NTS, we were unable to verify the linguistic accuracy of the generated representations for videos not included in our reference dictionary. Thus, we are unable to determine whether the model captures meaningful sign distinctions or merely produces stable but incorrect representations. Consequently, the utility of these annotations for downstream tasks such as keyword identification may rely more on internal consistency than on true linguistic validity, and the results should therefore be interpreted with caution.

8. Acknowledgements

We gratefully acknowledge *Tegnbanken* and the *SignPuddle* project for their valuable sign language resources, which were essential for our research on NTS.

9. Bibliographical References

Sobhan Asasi, Mohamed Ilyes Lakhal, Ozge Mercanoglu Sincan, and Richard Bowden. 2025. Beyond gloss: A hand-centric framework for gloss-free sign language translation. In *36th British Machine Vision Conference 2025, BMVC 2025, Sheffield, UK, November 24-27, 2025*. BMVA.

Félix Bigand, Elise Prigent, Bastien Berret, and Annelies Braffort. 2021. Machine learning of motion statistics reveals the kinematic signature of the identity of a person in sign language. *Frontiers in Bioengineering and Biotechnology*, 9:710132.

Danielle Bragg, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreault, 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, page 16–31, New York, NY, USA. Association for Computing Machinery.

Danielle Bragg, Oscar Koller, Naomi Caselli, and William Thies. 2020. Exploring collection of sign language datasets: Privacy, participation, and model performance. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '20, New York, NY, USA. Association for Computing Machinery.

Necati Cihan Camgöz, Oscar Koller, Simon Hadfield, and Richard Bowden. 2020. Sign language transformers: Joint end-to-end sign language recognition and translation. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10020–10030.

Amanda Duarte, Shruti Palaskar, Lucas Ventura, Deepti Ghadiyaram, Kenneth DeHaan, Florian Metze, Jordi Torres, and Xavier Giro-i Nieto. 2021. How2Sign: A large-scale multimodal dataset for continuous American Sign Language. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2734–2743.

Lindsay Nicole Ferrara, Benjamin Donald Anible, and Lena Mei Kalvenes Anda. 2023. Exploring sign-writing contact and multilingualism in the norwegian deaf community. In Ella Wehrmeyer, editor, *Advances in Sign Language Corpus Linguistics*, pages 66–89. John Benjamins Publishing Company.

Manfred Georg, Garrett Tanzer, Esha Uboweja, Saad Hassan, Maximus Shengelia, Sam Sepah, Sean Forbes, and Thad Starner. 2025. FSboard: Over 3 million characters of ASL fingerspelling collected via smartphones. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 13897–13906.

Rolf Piene Halvorsen. 2012. *Tre diskursmarkører i norsk tegnspråk: en studie av blunk, blikkendring og nikk i åtte fortellinger*. Ph.D. thesis, University of Oslo.

Rolf Piene Halvorsen. 2025. Norsk tegnspråk. In *Store Norske Leksikon*. Accessed: 2026-02-04.

[34] https://github.com/google-ai-edge/mediapipe/blob/master/docs/solutions/holistic.md

Thomas Hanke. 2004. HamNoSys – representing sign language data in language resources and language processing contexts. In *Proceedings of the LREC2004 Workshop on the Representation and Processing of Sign Languages: From SignWriting to Image Processing. Information techniques and their implications for teaching, documentation and communication*, pages 1–6, Lisbon, Portugal. European Language Resources Association (ELRA).

Håkon Bygdås Høgset. 2024. Machine learning approaches to automated sign language translation. Master's thesis, NTNU, Trondheim. Accessed: 2026-02-04.

Tommi Jantunen, Johanna Mesch, and Lindsay Ferrara. 2024. Corpus research on signed languages in the nordic countries. *Nordic Journal of Linguistics*, page 1–14.

Oscar Koller. 2020. Quantitative survey of the state of the art in sign language recognition. arXiv: 2008.09918.

Maria Kopf, Marc Schulder, and Thomas Hanke. 2022. The Sign Language Dataset Compendium: Creating an overview of digital linguistic resources. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources*, pages 102–109, Marseille, France. European Language Resources Association (ELRA).

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, Rotem Zilberman, and Ohad Langer. 2024. signwriting-evaluation: Effective sign language evaluation via signwriting. arXiv: 2410.13668.

Adrián Núñez-Marcos, Olatz Perez de Viñaspre, and Gorka Labaka. 2023. A survey on sign language machine translation. *Expert Systems with Applications*, 213:118993.

Justin M. Power, David Quinto-Pozos, and Danny Law. 2025. Advances in the historical linguistics of signed languages. *Language and Linguistics Compass*, 19(6):e70026.

Valerie Sutton. 1990. *Lessons in Sign Writing*. SignWriting.

Benjamin Svendsen and Seifedine Kadry. 2023. Comparative analysis of image classification models for norwegian sign language recognition. *Technologies*, 11(4).

Garrett Tanzer. 2025. Fingerspelling within sign language translation. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 385–464, Albuquerque, New Mexico. Association for Computational Linguistics.

Md. Zia Uddin, Costas Boletsis, and Pål Rudshavn. 2025. Real-time norwegian sign language recognition using mediapipe and lstm. *Multimodal Technologies and Interaction*, 9(3).

Zhaoyang Xia, Yang Zhou, Ligong Han, Carol Neidle, and Dimitris N. Metaxas. 2024. Diffusion models for sign language video anonymization. In *Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources*, pages 395–407, Torino, Italia. ELRA and ICCL.