Improving Sign Language machine translation results by using linguistic features and other data augmentation methods: *A series of studies on Spanish Sign Language (LSE) and German Sign Language (DGS)*

Euan McGill Horacio Saggion Santiago Egea Gómez Luis Chiruzzo

Universitat Pompeu Fabra, Barcelona

Universidad de la República de Uruguay

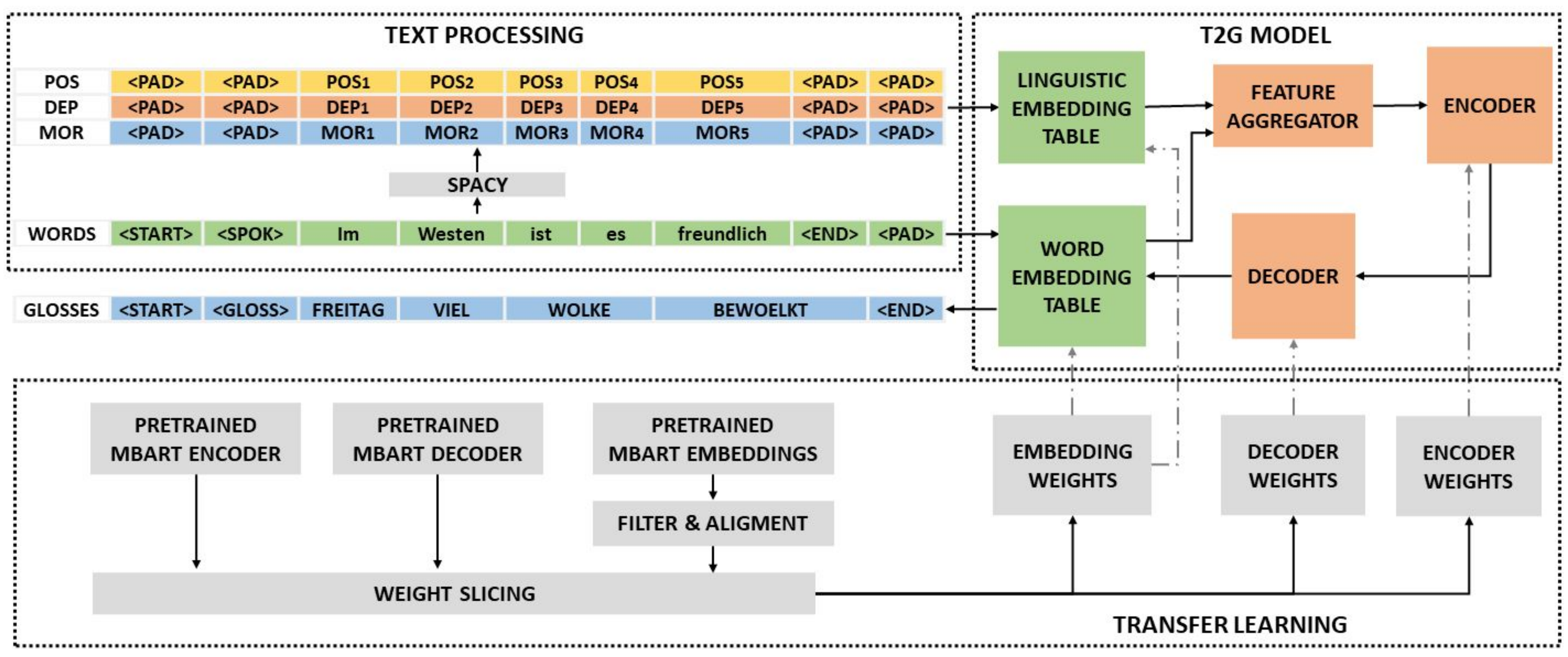
Background

There are approximately **466M**[1] people worldwide who experience hearing loss of some kind, of whom **70M**[2] communicate predominantly through one of **400+** attested[3] Sign Languages (SLs). SLs are fully-fledged languages with their **own complex grammar**, and are not typologically related to the spoken languages which coexist in the areas where they are used.

SLMT, or machine translation for signed languages, is a nascent technology, existing for little over 20 years compared with MT for spoken languages which has been studied since the 1950s. There are a few reasons for this:
(1) d/Deaf and Hard-of-Hearing (**DHH**) communities tend to be marginalised in society - with SLs only becoming officially recognised in European countries in the last few years.
(2) SLs are produced in the **visual-spatial modality** and in a **non-linear** manner where signs may co-occur in time and space and some of these grammatical features are still not fully understood by linguists.
(3) There is a **lack of writing system** which is widely used in SLs, so we often use **glosses** which provide a semantically lacking lexeme-based representation.
(4) The **data problem** - parallel datasets for SLMT are severely limited (De Sisto et al., 2022) in size, scope (SLMT is inherently **multimodal** between visual and textual representations (Bragg et al., 2019)), and standardisation of transcription methods (Cormier et al., 2016).

Our work intends to improve **text-to-gloss** and **gloss-to-text** translation, where glosses act as the **intermediate step** between SL video and spoken language text in the translation pipeline. End-to-end systems exist (e.g. Camgöz et al., 2020), but they are not yet optimally performing for all SL data, particularly that which is *extremely* low resource (Moryossef et al., 2021). We **inject linguistic features** into neural encoder-decoder translation models. This method is known as the **'factored transformer'** (Sennrich & Haddow, 2016) and has shown particular promise in improving low-resource MT results (Armegnol-Estapé et al., 2021).

For LSE, we also expand small SL datasets using a **rule-based transformation** to monolingual spoken language corpora into **synthetic glosses** based on the **grammatical rules** of a given SL.

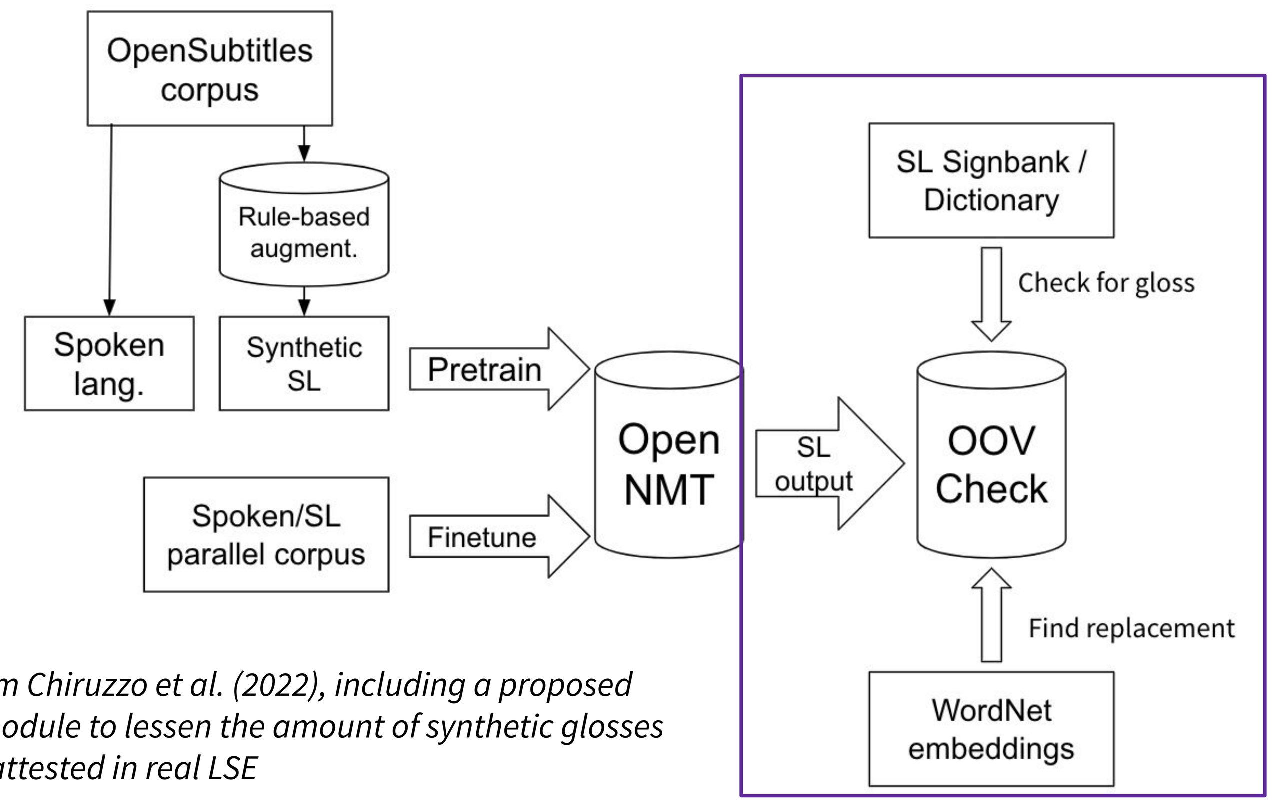


From Egea Gómez et al. (2022)

German→DGS Text2Gloss with linguistic features

Our first experiments in this area (Egea Gómez et al., 2021; 2022) involved enriching gloss representations with syntactic **dependency**, part-of-speech (**PoS**), and **morphological** information (model shown above). We used the ubiquitous PHOENIX-Weather corpus for comparability with other Text2Gloss studies.

Our 2022 study also experimented with different **feature aggregation** methods, or ways of including these linguistic features in the model. We found that using **convolutional block aggregation**, as well **transfer learning** of pre-trained model weights (rather than random initialisation) substantially improved translation performance. Our **best sacreBLEU score was 20.57**, comparable with the highest E2E strategies available at the time.

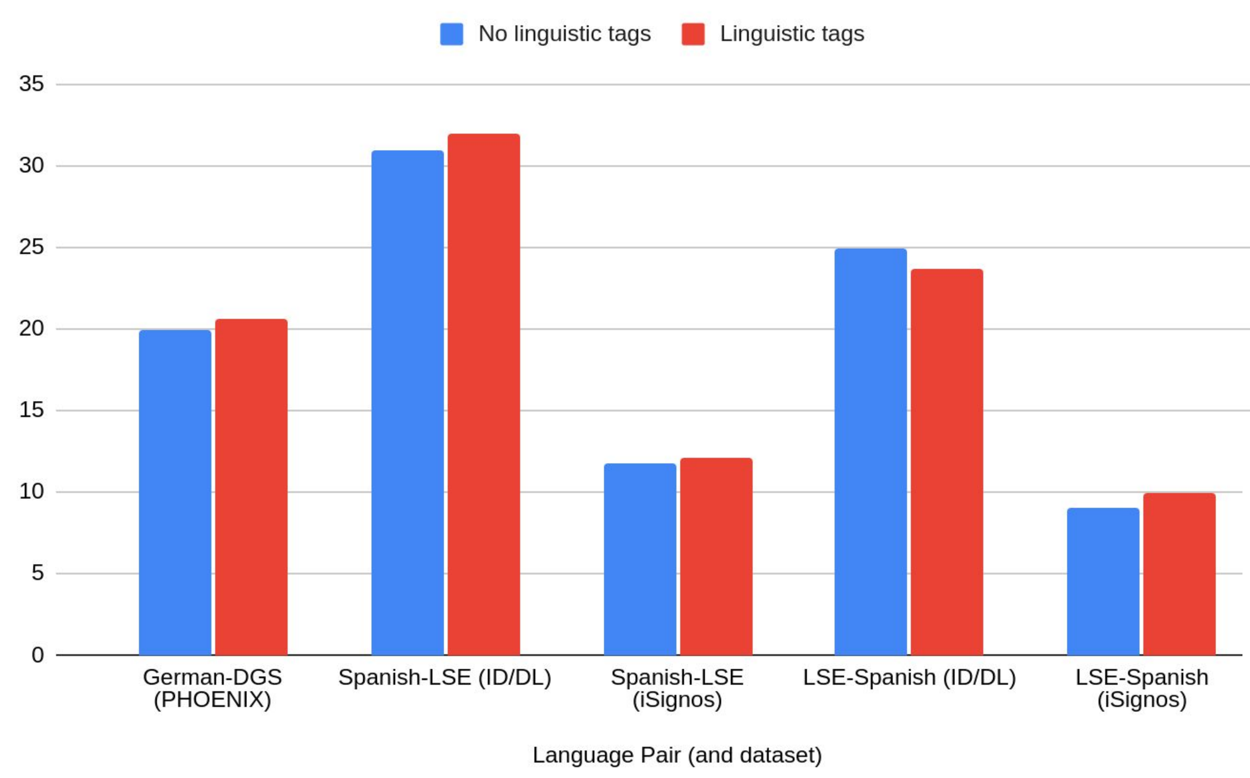


Model dgm from Chiruzzo et al. (2022), including a proposed OOV-checker module to lessen the amount of synthetic glosses which are not attested in real LSE

Spanish↔LSE Text2Gloss and Gloss2Text with linguistic features, pretrained with synthetic gloss data

We first **pre-trained** our model with parallel Spanish-*synthetic* LSE gloss data which we created with a **rule-based** strategy before fine-tuning on Spanish-LSE two different parallel corpora in each study respectively (Chiruzzo et al., 2022; McGill et al., 2023). We again found that linguistic features improved translation performance, but **only Text2Gloss** - this is likely because we used Spanish features for the unrelated language LSE, as there is **no NLP support** for LSE.

Our 2023 work manually PoS-tags LSE data, opening the door for **SL NLP** (Yin et al., 2021) techniques such as creating UD treebanks in future. When using these tags, we now saw improvements in translations in **both Gloss2Text** and Text2Gloss. We plan to do more work in future on automatically tagging SL data.



Summary of experimental results across our studies

BLEU Score	Interpretation
< 10	Almost useless
10 - 19	Hard to get the gist
20 - 29	The gist is clear, but has significant grammatical errors
30 - 40	Understandable to good translations
40 - 50	High quality translations
50 - 60	Very high quality, adequate, and fluent translations
> 60	Quality often better than human

BLEU interpretability - source: Google research

List of mentioned work:

- McGill, E., Chiruzzo, L., Egea Gómez, S., & Saggion, H. (2023) "Part-of-Speech tagging Spanish Sign Language data and its applications in Sign Language machine translation" *RESOURCEFUL@NoDaLiDa*
- Chiruzzo, L., McGill, E., Egea Gómez, S., & Saggion, H. (2022) "Translating Spanish into Spanish Sign Language: Combining Rules and Data-driven Approaches" *LoResMT@COLING*
- Egea Gómez, S., Chiruzzo, L., McGill, E., & Saggion, H. (2022) "Linguistically Enhanced Text to Sign Gloss Machine Translation" *NLDB@ACM*
- Egea Gómez, S., McGill, E. & Saggion, H. (2021) "Syntax-aware Transformers for Neural Machine Translation" *BUCC@RANLP*

Key references:

- Kearsy Cormier et al. (2016) "Digging into Signs: Emerging Annotation Standards for Sign Language Corpora"
- Rico Sennrich & Barry Haddow (2016) "Linguistic Input Features Improve Neural Machine Translation"
- Danielle Bragg et al. (2019) "Sign Language Recognition, Generation and Translation: An Interdisciplinary Perspective"
- Necati Cihan Camgöz et al. (2020) "Sign Language Transformers: Joint End-to-end Sign Language Recognition and Translation"
- Jordi Armegnol-Estapé et al. (2021) " Enriching the Transformer with Linguistic Factors for Low-Resource Machine Translation."
- Amit Moryossef et al. (2021) "Data Augmentation for Sign Language Gloss Translation"
- Kayo Yin et al. (2021) "Including Signed Languages in Natural Language Processing"
- Mirella De Sisto et al. (2022) "Challenges with Sign Language Datasets for Sign Language Recognition and Translation"

1) World Health Organisation
2) World Federation of the Deaf
3) SIL International

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017255

EXCELENCIA MARÍA DE MAEZTU