Sign Language Translation: Ongoing Development, Challenges and Innovations in the SignON Project

Dimitar Shterionov[*], Mirella De Sisto[*], Vincent Vandeghinste[†],, Aoife Brady[‡], Mathieu De Coster[§], Lorraine Leeson[¶], Josep Blat[**], Frankie Picron[††], Marcello Paolo Scipioni[‡‡], Aditya Parikh[§§], Louis ten Bosch[§§], John O'Flaherty[‖], Joni Dambre[§], Jorn Rijckaert[ˣ]

[*]Tilburg University, [†]Instituut voor de Nederlandse Taal, [‡]ADAPT, [§]Ghent University
[¶]Trinity College Dublin,[**]Universitat Pompeu Fabra,[††]European Union of the Deaf,
[‡‡]Fincons,[§§]Radboud University,[‖]mac.ie,[ˣ]Vlaams Gebarentaalcentrum

1 Introduction

SignON[1] focuses on the research and development of a *sign language (SL) translation mobile application and an open communications framework*. SignON addresses the lack of technology and services for the automatic translation between signed and spoken languages, through an inclusive, human-centric solution which facilitates communication between deaf, hard of hearing (DHH) and hearing individuals.

We present an overview of the status of the project, describing the milestones and the approaches developed to address the challenges and peculiarities of SL machine translation (SLMT).

SLs are the primary means of communication for over 70 million DHH individuals.[2] Despite this, they are rarely included in ongoing developments of natural-language processing (NLP) advancements (Yin et al., 2021). Machine translation (MT) research which targets SLs is still in its infancy, due mainly to the lack of data and effective representation of signs (including the lack of a standardized written form for SLs).

Both the low volume of available resources, as well as the linguistic properties of SLs provide challenges for MT. Furthermore, SLs are visual languages, which presents yet another challenge: the recognition and synthesis of a signing human.

2 The SignON approach to SLMT

The objective of the SignON project is MT between signed and spoken languages in all possible combinations, as well as the delivery of this service to the primary user groups: DHH and hearing users.

The project revolves around 4 spoken (English, Spanish, Dutch, Irish) and 5 SLs, (ISL, NGT, VGT, LSE, and BSL —namely Irish, Dutch, Flemish, Spanish and British SL). Addressing this many language pairs and directions on a pair-by-pair basis would require a substantial amount of time and effort, far beyond the scope of the project. SignON employs an MT approach that (i) focuses on processing and understanding individual languages, (ii) employs a common multi-lingual representation (InterL) to facilitate translation and (iii) uses symbolic as well as deep-learning methods for the synthesis of a 3D virtual signer. This approach involves automatic SL and speech recognition (SLR and ASR respectively), NLP, sign and speech synthesis, text generation and, most importantly, representation of utterances in a common frame of reference —an interlingual representation space based on embeddings and/or symbolic structures, the InterL. The complexity and diversity of these processing steps require multi-domain knowledge and expertise. Furthermore, we chose this approach as there are only limited parallel resources available between signed and spoken/written languages. Relying on techniques such as transfer learning, and pre-built NLP models (i.e. mBART (Lewis et al., 2020)) will improve MT performance.

We have built state-of-the-art models and components for SLR , exploiting convolutional

[1]SignON is a Horizon 2020 (Research and Innovation Programme Grant Agreement No. 101017255) project that runs from 2021 until the end of 2023. `https://signon-project.eu/`. The consortium is constituted by 17 partners, among which Instituut voor de Nederlandse Taal, Tilburg University, ADAPT, Ghent University, Trinity College Dublin, Universitat Pompeu Fabra, European Union of the Deaf, Fincons, Radboud University, mac.ie, Vlaams Gebarentaalcentrum, and Dublin City University.

[2]According to the World Federation of the Deaf.

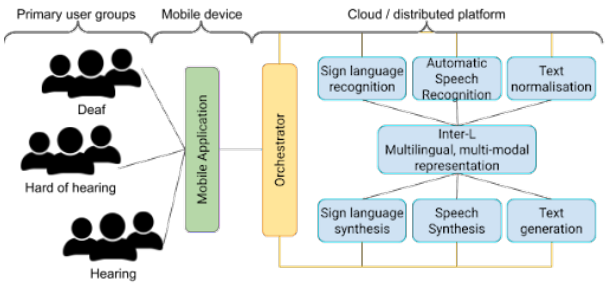

Figure 1: General approach of the SignON translation system

neural network-, recurrent neural network- and transformer-based models, natural-language understanding and MT based on mBART. We are developing approaches through wordnets and abstract semantic representation and synthesis based on language specific logical structures for SL, behavioral markup language and a 3D avatar rendering system.

The ASR component will tune to the use cases and to the speaker (including atypical speech from deaf speakers and speakers with cochlear implants). The ASR addresses (i) privacy challenges (ii) adaption to communicative settings and (iii) extension to new data and languages. Currently, English and Dutch are ready; Spanish is in progress. The transfer learning approach is adapted for Irish. The ASR works as a web service via a secure restful API.

3 SignON application and open framework

The general architecture (Figure 1) consists of a mobile application which connects users to the cloud-based MT platform. The SignON app is the interface between the user and the SignON framework which handles the internal data flow and processing. The *framework* executes the following steps. The source message (audio, video or text) and any relevant metadata coming from *the mobile app* is processed by an *orchestrator* which queues it towards the translation pipeline through a *message broker*. A *dispatcher* subscribed to the appropriate queue receives the message, invoking the relevant component depending on the type of input. After the required processing is complete, the message passes to the next stage of the pipeline until, finally, once the translation tasks are completed, the output message is produced in the requested format (text, audio or sign language avatar). The output is delivered to the app via the *orchestrator*. Each component is encapsulated in a docker container and distributed over different machines.

The first release of the SignON mobile application is due in June 2022, and will then evolve to its final release at the end of the project (Dec. 2023). The app will be available as open source and for free.

4 Societal impact

Along with the technological and academic innovations that come in terms of new models and methodsfor SLMT, SignON strives towards having a large societal impact. Currently we face societal challenges such as clashes between the views of DHH and hearing people, with respect to use-cases, technological importance and communication needs. We organized two sets of interviews with deaf participants, an online survey and we have two round tables planned. Via workshops we inform both the research and user communities about the progress of SignON and the state-of-the-art in SLMT.

5 Progress and next steps

In the first 15 months of this project 8 academic papers were accepted for publication. These papers discuss SLR, NLP, SLMT as well as SL representations. At the time of writing more than 5 papers are under review. We have conducted focus group interviews with VGT, ISL, LSE and NGT signers as well as public and internal surveys.

References

Lewis, Mike, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In D. Jurafsky et al., editor, *Proc. of the 58th Annual Meeting of the Assoc. for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 7871–7880. ACL.

Yin, Kayo, Amit Moryossef, Julie Hochgesang, Yoav Goldberg, and Malihe Alikhani. 2021. Including signed languages in natural language processing. In *Proc. of the 59th Annual Meeting of the ACL and the 11th Int. Joint Conference on NLP (Volume 1: Long Papers)*, pages 7347–7360, Online, August. ACL.